RICE UNIVERSITY

Three Essays on National Oil Company Efficiency, Energy Demand and Transportation

by

Stacy L. Eller

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Peter R. Hartley, George and Cynthia Mitchell Chair in Sustainable Development and Environmental Economics, Professor of Economics

Peter Mieszkowski, Professor Emeritus

Kenneth B. Medlock III, James A. Baker, III, and Susan G. Baker Fellow in Energy and Resource Economics

HOUSTON, TEXAS
DECEMBER 2010

ABSTRACT

Three Essays on National Oil Company Efficiency,

Energy Demand and Transportation

by

Stacy L. Eller

This dissertation is composed of three separate essays in the field of energy economics. In the first paper, both data envelopment analysis and stochastic production frontier estimation are employed to provide empirical evidence on the revenue efficiency of national oil companies (NOCs) and private international oil companies (IOCs). Using a panel of 80 oil producing firms, the analysis suggests that NOCs are generally less efficient at generating revenue from a given resource base than IOCs, with some exceptions. Due to differing firm objectives, however, structural and institutional features may help explain much of the inefficiency.

The second paper analyzes the relationship between economic development and the demand for energy. Energy consumption is modeled using panel data from 1990 to 2004 for 50 countries spanning all levels of development. We find the relationship between energy consumption and economic development corresponds to the structure of aggregate output and the nature of derived demand for electricity and direct-use fuels in each sector. Notably, the evidence of non-constant income elasticity of demand is much greater for electricity demand than for direct-use fuel consumption. In addition, we show that during periods of rapid economic development, one in which the short-term growth

rate exceeds the long-run average, an increase in aggregate output is met by less energy-efficient capital. This is a result of capital being fixed in the short-term. As additional, more efficient capital stock is added to the production process, the short-term increase in energy intensity will diminish.

In the third essay, we develop a system of equations to estimate a model of motor vehicle fuel consumption, vehicle miles traveled and implied fuel efficiency for the 67 counties of the State of Florida from 2001 to 2008. This procedure allows us to decompose the factors of fuel demand into elasticities of vehicle driving demand and fuel efficiency. Particular attention is paid to the influence of the price of fuel, the sale of goods and services, vehicle ownership and population density on each component of our model.

ACKNOWLEDGEMENTS

I would like to express my most sincere gratitude to the members of my committee for their endless encouragement and inspiration. First, I thank Dr. Peter R. Hartley, the chair of my committee, for his advice and guidance throughout my graduate school career. His passion for teaching never ceases to amaze me. I also want to thank Dr. Peter Mieszkowski. After several difficult semesters, his course in public economics was a turning point for me. The subject matter was fascinating, and with his encouragement, I developed the confidence needed to continue down the long road toward my defense. Last but not least, I want to thank the third member of my committee, Dr. Kenneth B. Medlock III. Ken was a colleague of mine at El Paso Merchant Energy, and this journey began almost a decade ago at his suggestion. Over the years, Ken has been a boss, an advisor, a mentor and a friend. With Ken's support I have achieved more than I would have ever imagined.

I must express a special thanks to Amy Jaffee and the Baker Institute for financial support, Professors Robin Sickles and Ronald Soligo for years of advice, and Altha Rodgers for her kindness and care. In addition, I thank fellow graduate students Timothy Gunning, Jennifer Rosthal, David Blazek and Ozgur Inal for their friendship.

Finally, I want to thank my family. The love and encouragement of my parents, Butch and Clara Eller, cannot be overstated. Without their support, none of this would have been possible. I thank my brother, Glenn, for teaching me the true value of competition. Lastly, I want to express sincere gratitude to my grandparents, Walt and Nancy Eller as well as Clifford and Sue Rackley.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. Essay 1 – Empirical Evidence on the Operational Efficiency of National Oil Companies[1]

a. Introduction

Countries that possess a national oil company (NOC) have held nearly 90% of global conventional crude oil reserves in every year since 1991, according to the U.S. Energy Information Agency (EIA, reporting data from *Oil & Gas Journal*). In addition, crude oil production in these countries increased from 50% of the global total in the mid-1980s to remain near 66% since the mid-1990s. Moreover, private international oil companies (IOCs) have found it increasingly difficult to find or gain access to large oil and gas prospects, which could further increase NOC dominance of the global crude oil market.

In a recent paper, Hartley and Medlock (2007) developed and analyzed a theoretical model of the operation and development of a National Oil Company (NOC). They argued that the political overseers of a NOC are likely to require it to pursue various non-commercial objectives. Accordingly, a NOC is likely to favor excessive employment while under-investing in reserves and shifting resource extraction away from the future toward the present. A NOC also may be forced to sell oil products to domestic consumers at subsidized prices. Thus, NOCs will generally appear to be relatively inefficient, especially at generating revenue from given inputs of labor and reserves, when compared to private IOCs.[2] If the analysis is correct, future global crude oil prices are likely to be higher than they would be if NOCs were less dominant.

[1] This is a joint paper with Peter R. Hartley, and Kenneth B. Medlock III

[2] The word "inefficient" is used to mean getting less of a specified output from a given input bundle. The measures we calculate need not correspond to the economic notion of "efficiency" as Pareto optimality. In

In this paper, we apply both non-parametric data envelopment analysis (DEA) and a parametric stochastic frontier analysis (SFA) to a sample of 80 firms over three years (2002 to 2004) to assess whether NOCs are indeed relatively revenue inefficient. We examine *revenue* efficiency not only because revenue is a key objective for both public and private firms, but also because political pressure is likely to force a NOC to sell products to domestic consumers at subsidized prices. Physical output measures would not necessarily capture the effect of such subsidies.

A substantial body of evidence shows that state-owned firms generally are less efficiently managed than their private sector counterparts.[3] Recent theoretical analyses of relative firm efficiency have focused on the principal-agent paradigm.[4] Specifically, if managers (agents) seek to maximize their own utility rather than the objectives of the owners (principals) firm inefficiency can result. Theory and evidence suggest, however, that institutional features of private ownership help align managerial objectives with those of the owner. For example, tradable ownership shares give owners an incentive to monitor managerial performance and provide them with a readily observable measure of managerial performance. Furthermore, poor performance reduces the price of ownership shares, increasing the threat of takeover and decreasing the manager's job security. The threat of bankruptcy, which is generally absent for state-owned enterprises with government guaranteed debt, can also encourage managers to maintain cash flows. Such institutional features provide some degree of accountability and thus improve the

---

particular, a NOC maximizing an objective function in which revenue is only one argument could be economically efficient. Nevertheless, the extent to which firms generate different amounts of revenue from a given vector of inputs can be used to judge whether their objectives truly differ as hypothesized.

[3] Villalonga (2000) provides a thorough discussion of theoretical and empirical studies of the efficiency of publicly versus privately owned firms, as well as the efficiency implications of privatization.

[4] The principal-agent paradigm was introduced by Jensen and Meckling (1976) and Harris and Raviv (1978). Property rights theory was originally developed by Alchian (1965). Other notable work on privatization includes Laffont and Tirole (1991) and Schmidt (1996).

efficiency of private firms. Conversely, their absence in state-owned firms can help explain relative inefficiency.

More to the point, while the owners of a private firm all desire the managers to maximize the firm's market value, the political overseers of managers of state-owned enterprises generally have a less well-defined objective. Policies that engender excessive employment and subsidized consumer prices can increase political support. Regardless of their success in achieving a political objective, such policies hinder firm efficiency.

There is an extensive literature on estimating productive efficiency. The two primary methods for computing production frontiers are data envelopment analysis and stochastic frontier analysis. In summary, DEA uses linear programming techniques to construct a non-parametric piecewise-linear convex hull of observed input-output bundles. In contrast, SFA involves estimating the production frontier from observed input-output bundles using econometric methods. By definition, the observed input-output bundle of an efficient firm is on the production frontier, where as an inefficient firm is off the production frontier.

Despite the importance of NOCs in the world oil market, we know of only one article, by Al-Obaidan and Scully (1991), that examines the relative efficiencies of NOC's using parametric techniques such as stochastic frontier analysis, and we could not find any studies using non-parametric data envelopment analysis. One potential explanation for a lack of substantial analysis of NOC efficiency is the paucity of data on non-publicly traded firms. Regardless of the reason, the Al-Obaidan and Scully paper is relatively novel in its application.

Using data for 44 firms in a single year, 1981, Al-Obaidan and Scully construct a production frontier using deterministic and stochastic methods to examine the ability of firms to use assets and employees to produce output, where output was defined as either revenue earned or the quantity of crude oil produced plus the quantity of crude oil processed. Relative to private firms, the authors found that NOCs are only 63% to 65% as efficient in generating revenue.

Although our results are generally consistent with those of Al-Obaidan and Scully, our study differs in many respects. Most differences involve the data used in the analysis. For example, Al-Obaidan and Scully omit all OPEC nations, arguing that the demonstrated efficiency of those firms is "related more to the accident of geography than to the allocation of resources within the firm." In addition, they considered only vertically integrated firms, omitting firms specializing in the downstream or upstream sectors. In contrast, we omit only firms that specialize in the downstream sector.[5] Another key difference is that we use panel data in our analysis.

Because the merits of the two very different approaches are often debated, we present results from both DEA and SFA to highlight the fact that evidence of NOC inefficiency is robust to the technique used to measure efficiency. Therefore, we are able to capitalize on the strengths, and guard against the weaknesses, of each method. In the following section, we describe DEA and SFA in greater detail. Section b summarizes the data we employ, while section c presents and discusses the basic results. Section d expands on section c by explaining the observed firm inefficiency relative to institution and structural features of specific firms. We follow with some concluding remarks.

Estimating the Efficiency of NOCs

---

[5] We omitted such firms because we consider reserves as one of the inputs.

Koopmans (1951) defined a producer as *technically* efficient if and only if increasing production of one output results in either a reduction of some other output or an increase in some input. Debreu (1951) and Shephard (1953) suggested using distance functions to measure a firm's technical efficiency. Debreu's output-oriented measure requires a firm to maximize output for a given set of inputs, while in Shephard's input-oriented definition a technically efficient firm minimizes the use of inputs for a given output. These measures are equivalent when the production technology exhibits constant returns to scale (CRS).

Farrell (1957) was the first to measure technical efficiency by applying Debreu's measure to the U.S. agricultural sector. If a firm's observed output for a given level of inputs is best in practice, the firm is said to be *on* the frontier. In order to identify best practice, Farrell suggested constructing a piecewise-linear convex hull of observed input-output bundles. Authors such as Boles (1966) used non-parametric mathematical programming, later termed data envelopment analysis (DEA) by Charnes, Cooper, and Rhodes (1978), to identify such a piecewise-linear convex hull.

More specifically, the output-oriented DEA technique calculates the proportion to which a firm maximizes output (physical units of production or monetary units of revenue or profit[6]) for a given level of inputs on a scale from zero to one. Specifically, suppose we have data for N firms each using K inputs and producing M outputs. Define X as the $K \times N$ matrix of inputs, Y as the $M \times N$ matrix of outputs, and let x_n and y_n denote the inputs and outputs, respectively, of firm *n*. The CRS output-oriented technical efficiency of each firm is then calculated by solving the following linear program:

---

[6] We will focus on *revenue* efficiency by specifying the firm's output to be revenue.

$$
\begin{aligned}
& \max_{\lambda,\theta} \theta \\
\text{subject to:} \quad & -\theta y_n + Y\lambda \geq 0 \\
& x_n - X\lambda \geq 0 \\
& \lambda \geq 0
\end{aligned}
\tag{I-1}
$$

where $1 \leq \theta \leq \infty$ is a scalar and λ is an $N \times 1$ vector of constants. The technical efficiency score is then defined as θ^{-1}.[7]

Farrell's research in productive efficiency also influenced the development of stochastic frontier analysis, the other method employed in this paper to estimate firm-specific measures of revenue efficiency. Stochastic frontier analysis was introduced simultaneously by Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977).[8] Unlike the DEA approach, stochastic frontier analysis requires various parametric assumptions. The SFA approach identifies inefficiency as part of a two-component error, where one component captures statistical noise and the other captures inefficiency.

Specifically, consider a single output production function for a cross-section of N firms with K inputs to be given as $y_n = f(x_{1,n}, \ldots, x_{K,n})$. If we assume the production technology can be represented as Cobb-Douglas, then we can linearize the production function by taking the natural logarithm, yielding

$$
\ln y_n = \alpha + \sum_{k=1}^{K} \beta_k \ln x_{k,n} + v_n - u_n \tag{I-2}
$$

[7] Banker, Charnes and Cooper (1984) extended the CRS DEA model to allow variable returns to scale (VRS). Coelli, Rao, and Battese (1999) provide a thorough discussion of DEA for both CRS and VRS.
[8] Prior to the development of the stochastic frontier approach, authors such as Aigner and Chu (1968) and Afriat (1972) estimated a *deterministic* production frontier from a cross-section of firms. The resulting one-sided, nonnegative residual from that method captures each firm's inefficiency. Such a deterministic approach, however, does not allow random factors, such as unusual weather, to influence output.

where v_n is a stochastic component generally assumed to be normally distributed and u_n captures the nonnegative technical efficiency component. Define ε_n as the composed error such that $\varepsilon_n = v_n - u_n$. Jondrow et al. (1982) proposed estimating individual technical efficiency from the expected value of u_n conditional on ε_n using maximum likelihood. Firm-specific efficiency can then be calculated as $e^{-E(u_n|\varepsilon_n)}$. Since $u_n \geq 0$, efficiency will be bounded between zero and one.

Kumbhakar (1987) and Battese and Coelli (1988) proposed a panel data extension to this maximum likelihood approach to estimate firm-specific technical efficiency.[9] The log-linear Cobb-Douglas production function for panel data is

$$\ln y_{n,t} = \alpha + \sum_{k=1}^{K} \beta_k \ln x_{k,n,t} + v_{n,t} - u_n \tag{I-3}$$

where it is assumed the technical efficiency is time-invariant.[10] Estimation of this production function using maximum likelihood requires the following distributional assumptions on the error components, $v_{n,t}$ and u_n:

(i) $v_{n,t} \sim N(0, \sigma_v^2)$,

(ii) $u_n \sim N^+(\mu, \sigma_u^2)$ where N^+ denotes the truncated-normal distribution, and

(iii) $v_{n,t}$ and u_n are distributed independently of each other and the regressors.

---

[9] Schmidt and Sickles (1984) also proposed using one-sided fixed-effects and random-effects to measure time-invariant producer-specific technical efficiency. See Kumbhakar and Lovell (2000) for a thorough survey of panel stochastic frontier analysis.

[10] The time-invariant specification of the panel frontier estimation was chosen because our panel contains observations for only three years, 2002 through 2004. The time-varying specifications following Cornwell, Schmidt and Sickles (1990) and Battese and Coelli (1992), for example, are generally used for longer panels in which the focus of the research is to measure efficiency change over some period such as deregulation or privatization.

Given these distributional assumptions on the composed error, defined as $\varepsilon_{n,t} = v_{n,t} - u_n$, the log likelihood function for a sample of *N* producers over *T* time periods is

$$\begin{aligned}\ln L = \text{constant } & - \frac{N(T-1)}{2}\ln\sigma_v^2 - \frac{N}{2}\ln\left(\sigma_v^2 + T\sigma_u^2\right) - N\ln\left[1-\Phi\left(-\frac{\mu}{\sigma_u}\right)\right] \\ & + \sum_n \ln\left[1-\Phi\left(-\frac{\tilde{\mu}_n}{\tilde{\sigma}}\right)\right] - \frac{\sum_n \varepsilon_n{}'\varepsilon_n}{2\sigma_v^2} - \frac{N}{2}\left(\frac{\mu}{\sigma_u}\right)^2 + \frac{1}{2}\sum_n\left(\frac{\tilde{\mu}_n}{\tilde{\sigma}}\right)^2\end{aligned} \tag{I-4}$$

where $\Phi(\bullet)$ is the standard normal cumulative distribution function, and

$$\tilde{\mu}_n = \frac{\mu\sigma_v^2 - T\bar{\varepsilon}\sigma_u^2}{\sigma_v^2 + T\sigma_u^2}, \tilde{\sigma}^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma_v^2 + T\sigma_u^2}, \text{ and } \bar{\varepsilon} = \frac{1}{T}\sum_t \varepsilon_{n,t}.$$

This log likelihood function can be maximized to obtain parameter estimates. Furthermore, time invariant firm-specific estimates of technical efficiency can be recovered from the expected value of u_n conditional on $\varepsilon_{n,t}$. This is given by

$$E\left(u_n \mid \varepsilon_{n,t}\right) = \tilde{\mu}_n + \tilde{\sigma}\left[\frac{\phi(-\tilde{\mu}_n/\tilde{\sigma})}{1-\Phi(-\tilde{\mu}_n/\tilde{\sigma})}\right] \tag{I-5}$$

where $\phi(\bullet)$ and $\Phi(\bullet)$ are the standard normal density function and cumulative distribution function, respectively. The firm-specific technical efficiency score for firm *n* is equal to $e^{-E(u_n|\varepsilon_{n,t})}$, which is bounded between zero and one.

Both DEA and SFA have been used extensively to analyze productive efficiency. Comparison of methods is available in Gong and Sickles (1992), Banker (1993), Cooper and Tone (1997), and Ruggiero (2007), to name a few. However, we explore both approaches to highlight the fact that evidence of NOC inefficiency is robust to the approach used, rather than to compare methods for measuring efficiency.

The advantage of using the DEA approach is that it requires no assumptions regarding the functional form of the production technology and is not subject to the potential problems of assuming an inappropriate distribution of the error term. In this sense, the non-parametric DEA approach is more robust than the stochastic frontier method. However, since DEA does not account for statistical noise, estimates of efficiency will be biased when stochastic elements are a prominent feature of the true production process or the variables used in the analysis are measured with error.

The stochastic frontier approach, however, more directly shows how various factors influence firm behavior. It also allows more types of variables to be included in the analysis and provides a statistical measure of how well the proposed model explains the data. A potential weakness, however, is that one has to specify a structural relationship between inputs and outputs, including how stochastic terms arise. If these auxiliary assumptions are inaccurate, the resulting inferences about the underlying model may be compromised.

b. Data

We focus on the efficiency in generating *revenue* from inputs of employees, oil reserves, and natural gas reserves for 80 firms worldwide, including 10 of OPEC's 12 member nations.[11] Since OPEC's role in the international oil markets cannot be

[11] OPEC increased to 12 members when Angola officially joined in 2007. Although Angola is included in the study, Iraq is omitted due to ongoing domestic and petroleum industry turmoil and Libya due to a lack of relevant data.

overstated, the inclusion of these firms is important when estimating efficiency in the petroleum industry.

As noted in the introduction, we use revenue as a measure of output because, unlike output measured as a physical quantity, this allows us to capture the effects of forcing NOCs to subsidize domestic energy prices. In addition, we want a scalar measure of output for the stochastic frontier analysis and it is natural to aggregate different products using their relative market prices. The techniques used to measure technical efficiency can be directly applied to revenue efficiency. The interpretation, however, is slightly different because revenue efficiency requires a firm to be both technically efficient and to produce the appropriate mix of outputs given market prices.[12]

We use oil and gas reserves (measured separately) and total employment as inputs to revenue production. In contrast to Al-Obaidan and Scully, we do not include total assets as an input, primarily because data on total assets is not available for many NOCs, especially members of OPEC. By using total reserves as a measure of accumulated investment, we increase our sample by eight highly influential NOCs. Furthermore, reserves are likely to be a substantial part of total assets of most oil and gas firms, and are likely to be measured much more accurately than other assets.[13]

By ignoring non-reserve assets, however, we may be introducing another bias. Vertically integrated firms engaged in both upstream (exploration and production) and

[12] Kumbhakar and Lovell (2000) provide a detailed discussion of revenue, profit, and cost efficiency.

[13] For many oil and gas firms, reserves are audited and valued at market prices. By contrast, the remaining assets are valued at their book (or accounting) value rather than their economic value. Cumulative depreciation of non-reserves assets is an accounting measure correlated with the purchase price and age of the assets, but not necessarily with their productive capability. The accounting measure of asset value is also seriously distorted by inflation, which is important for many of the countries in our sample. Thus, the book value of total assets may be either over or understated relative to their economic value as an input to production. Consequently, using asset book value would impact the estimation of revenue efficiency in a way that would be difficult to interpret.

downstream (refining) operations will record revenue from the sale of refined products in addition to the external sale of crude to other parties. Since we are not measuring capital employed in refining, transporting and marketing operations as inputs, however, a vertically integrated firm would appear to be more efficient than other firms in generating revenue using the same reserves and employees.

The data were collected from the *Energy Intelligence* annual publication "Ranking the World's Oil Companies". Where available, company annual reports were used to verify the published data and to provide missing data. After eliminating firms specialized in downstream activities and thus lacking reserves, and firms for which other relevant data is unavailable, our sample was reduced from 100 to 80 firms (78 with complete information). Table I-1 lists the 80 firms in the study along with the countries in which their head offices are located, their shares of government ownership, and some statistics on revenue per employee and revenue per unit of reserves for 2004.[14]

We have also included in the table averages of revenue per employee and revenue per unit of reserves for different categories of firms. These indicate that the major international oil companies (BP, Chevron, ConocoPhillips, ExxonMobil, and Shell collectively denoted *major* IOCs) fall near the top of all the firms in the sample in both of the revenue per unit of input measures. In addition, although NOCs can be found throughout the range, on average they are toward the bottom of the table and dominate the bottom 20% in both measures.

[14] For this table, reserves are defined as the sum of crude oil reserves and natural gas reserves on a barrel of oil equivalent (boe) basis.

Table I-1: Companies with selected statistics for 2004

Company	Revenue per Employee *$1,000/employee*	Revenue per Reserves *$/boe*	Government Ownership *%*	Country
	NOCs			
Adnoc	205	0.20	100%	UAE
CNOOC	2,656	2.97	71%	China
Ecopetrol	824	2.26	100%	Colombia
Eni	1,056	10.50	30%	Italy
Gazprom	103	0.16	51%	Russia
INA	187	11.70	75%	Croatia
KMG	200	1.27	100%	Kazakhstan
KPC	1,650	0.34	100%	Kuwait
MOL	635	42.37	25%	Hungary
NIOC	283	0.11	100%	Iran
NNPC	1,460	0.56	100%	Nigeria
Norsk Hydro	673	11.37	44%	Norway
OMV	2,214	8.90	32%	Austria
ONGC	298	2.11	84%	India
PDO	1,591	0.98	60%	Oman
PDVSA	1,985	0.66	100%	Venezuela
Pemex	506	4.01	100%	Mexico
Pertamina	453	0.73	100%	Indonesia
Petrobras	773	3.39	32%	Brazil
PetroChina	111	2.52	90%	China
Petroecuador	1,026	1.51	100%	Ecuador
Petronas	1,202	1.45	100%	Malaysia
PTT	2,896	16.68	100%	Thailand
QP	1,800	0.10	100%	Qatar
Rosneft	86	0.19	100%	Russia
Saudi Aramco	2,261	0.40	100%	Saudi Arabia
Sinopec	192	19.76	57%	China
Socar	53	0.74	100%	Azerbaijan
Sonangol	755	1.37	100%	Angola
Sonatrach	688	0.93	100%	Algeria
SPC	375	1.71	100%	Syriac
Statoil	1,910	10.85	71%	Norway
TPAO	154	1.53	100%	Turkey
Average	*947.31*	*4.98*		

	Major IOCs			
BP	2,788	15.68	0%	UK
Chevron	2,606	12.78	0%	US
ConocoPhillips	3,368	14.03	0%	US
Exxon Mobil	3,148	12.26	0%	US
Shell	2,418	21.67	0%	Netherlands
Average	*2,865.48*	*15.28*		

Company	Revenue per Employee *$1,000/employee*	Revenue per Reserves *$/boe*	Government Ownership *%*	Country
		Others		
Amerada Hess		16.07	0%	US
Anadarko	1,838	2.52	0%	US
Apache	2,019	2.71	0%	US
BG	1,547	3.64	0%	UK
Burlington	2,537	2.74	0%	US
Chesapeake Energy	1,577	3.22	0%	US
CNR	4,606	3.85	0%	Canada
Devon	2,356	4.33	0%	US
Dominion	847	13.81	0%	US
EnCana	2,915	4.48	0%	Canada
EOG	1,844	2.38	0%	US
Forest Oil	1,841	4.02	0%	US
Husky Energy	2,149	9.53	0%	Canada
Imperial	2,838	17.91	0%	Canada
Kerr-McGee	1,263	4.15	0%	US
Lukoil	233	1.68	0%	Russia
Maersk	60	2.90	0%	Denmark
Marathon	1,757	39.14	0%	US
Murphy	1,436	21.60	0%	US
Newfield	2,114	4.45	0%	US
Nexen	1,048	4.25	0%	Canada
Nippon Oil	2,690	131.74	0%	Japan
Noble	2,433	2.54	0%	US
Novatek	220	0.21	0%	Russia
Occidental	1,577	4.46	0%	US
Penn West	1,577	2.53	0%	Canada
Petro-Canada	2,370	9.24	0%	Canada
PetroKazakhstan	546	4.12	0%	Kazakhstan
Pioneer	1,183	1.76	0%	US
Pogo	5,088	4.38	0%	US
Repsol YPF	1,561	10.79	0%	Spain
Santos	789	1.92	0%	Australia
Sibneft	189	1.81	0%	Russia
Suncor	1,447	13.41	0%	Canada
Surgutneftegas	121	1.01	0%	Russia
Talisman	2,207	3.26	0%	Canada
TNK	63	1.66	0%	Russia
Total	1,406	14.33	0%	France
Unocal	1,259	4.63	0%	US
Vintage	1,136	1.76	0%	US
Woodside	758	2.11	0%	Australia
XTO	1,437	1.94	0%	US
Average	*1,628.94*	*9.26*		

While the data in Table I-1 reveal that large firms with publicly held shares tend to have the highest revenue per unit of input measures, this may not mean they are more efficient. For example, these companies may be more vertically integrated than the remaining firms. The subsequent formal analysis will control for vertical integration by including a measure of petroleum product sales divided by total liquids production.[15]

The data suggest that a higher degree of government ownership may reduce efficiency in producing revenue from employees and reserves. For example, a government may force its NOC to subsidize domestic prices of oil products (a practice sometimes referred to as two-tiered pricing). In the formal analysis, we use information on the domestic retail prices of gasoline and diesel fuel to define a two-tiered pricing variable. While it is a rough approximation, we assume that every country with an average price below that of the United States has two-tier pricing.[16]

c. Analysis and Results

To begin, we used output-oriented DEA assuming CRS to calculate the firm-specific revenue efficiency measures using the observed input-revenue bundles for each year.[17] Employees, oil reserves (in million barrels), and natural gas reserves (in billion cubic feet) are included as separate inputs for the generation of revenue (in million U.S.

---

[15] This definition of vertical integration is consistent with the inverse of the integration ratio used by Al-Obaidan and Scully. Specifically, the authors define the integration ratio as barrels of oil produced divided by barrels oil processed. The implication of vertical integration, however, is the same.

[16] Metschies (2003) and Metschies (2005) provide a survey of international fuel prices for 2002 and 2004, respecitively. Following Metschies' definition of fuel subsidization, we assume that any country in which the simple average price of diesel and gasoline for 2002 and 2004 is below that of the U.S. subsidizes fuel prices.

[17] Calculations were performed using Coelli's software program *DEAP Version 2.1*. As the DEA model becomes more complex through the addition of structural and institutional variables in the following section, the efficiency scores for CRS converge to those of VRS. This suggests that the underlying technology truly exhibits CRS. Since all results and conclusions are identical for VRS, we report only CRS for simplicity.

dollars) for 78 firms from 2002 through 2004.[18] The average revenue efficiency score for each firm across the three years is graphed in the horizontal dimension in Figure 1.[19] To summarize, the IOCs are clustered near the frontier (right boundary), while the NOCs, although dispersed throughout, tend to be clustered near the left boundary. The average revenue efficiency measure using DEA for the NOCs is about 0.28 compared to 0.73 for the five major IOCs and 0.45 for other firms.

A second set of revenue efficiency measures was obtained using maximum likelihood to estimate a stochastic frontier for the 80 firms in our sample. We estimate a fairly simple model in which, as in the DEA approach, revenue is produced using oil reserves, natural gas reserves and employees as inputs. Time effects are included because the prevailing market prices of oil and gas, which are not constant across years, will affect the revenue generated in each year. Time effects were unnecessary in the DEA approach because separate revenue efficiency measures are calculated for each year. The equation estimated using the stochastic frontier approach is given as (standard errors in parentheses):

$$\begin{aligned} \ln Rev_{n,t} = {} & \underset{(0.5960)}{4.9421} + \underset{(0.0638)}{0.3786} * \ln L_{n,t} + \underset{(0.0612)}{0.1203} * \ln OilRsv_{n,t} + \underset{(0.0551)}{0.1888} * \ln NGRsv_{n,t} \\ & + \underset{(0.0247)}{0.2706} * t_{2003} + \underset{(0.0250)}{0.4416} * t_{2004} + v_{n,t} - u_n \end{aligned} \quad . \quad \text{(I-6)}$$

Each of the coefficients has the expected sign and is significantly different from zero at the 5% level.[20]

---

[18] We are limited to 78 firms here because two firms have an incomplete time series of data. Since we report average revenue efficiency scores for each year for DEA, using only the available data may be misleading. Thus we chose to drop these two firms, Socar and SPC.

[19] Averaging firm-specific efficiency over time using DEA is consistent with Gong and Sickles (1992), who use Monte Carlo techniques and panel data to compare stochastic frontier analysis and DEA.

[20] To compute firm-specific efficiency for panel data, DEA calculates a production frontier for each year which allows technology to change from one year to the next. In contrast, time-invariant stochastic frontier analysis with time effects, allows the intercept to shift from year to year, but the marginal productivity of inputs is assumed to be constant. Therefore, we test the assumption that the data can be pooled across time.

To determine whether revenue inefficiency does indeed exist in our sample, we test if $\gamma = \sigma_u /(\sigma_u + \sigma_v)$ is greater than zero. For this simple stochastic frontier model, γ is equal to 0.9593. Not only do we reject that γ is statistically greater than zero, γ is statistically greater than 0.5 which indicates that the composed error is dominated by the inefficiency term.[21]

As noted previously, the revenue efficiencies of firms as estimated in the stochastic frontier regression are given by $e^{-E(u_n|\varepsilon_{n,t})}$. Figure I-1 plots the resulting efficiencies in the vertical dimension for the 78 firms in our sample with complete data. Once again, we find the five major IOCs clustered near the frontier (upper boundary). Moreover, the NOCs, while dispersed, tend to be grouped near the lower boundary. Since the dominant tendency in Figure I-1 is for the observations to fall on a diagonal from the bottom left to the top right of the figure, the two efficiency measures are obviously positively correlated. The Spearman rank order correlation between the DEA and stochastic frontier efficiency measures is 0.6974 and a test of the null hypothesis that the two orderings are independent is rejected at a *p*-value less than 0.0001.

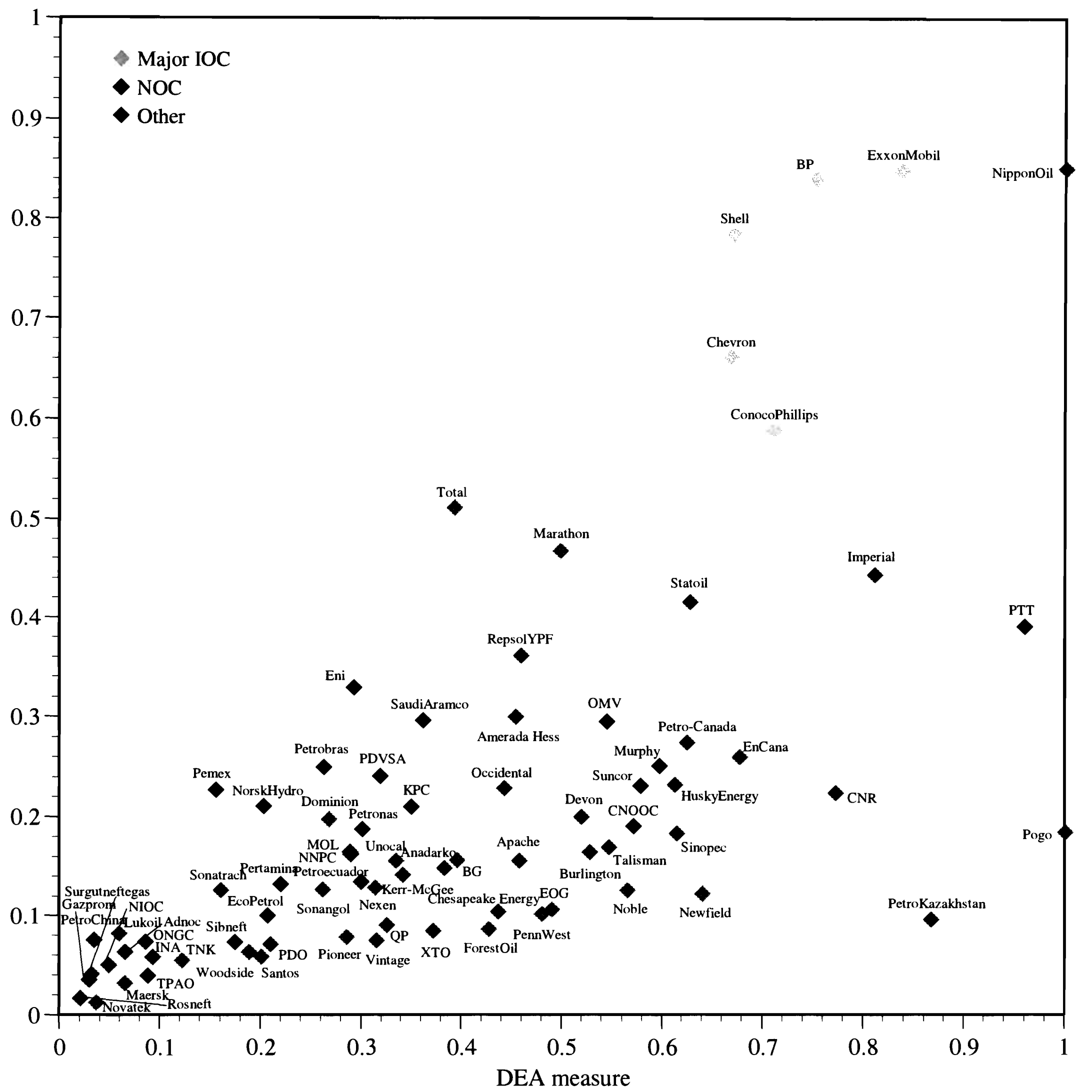
Similarly, a regression of the stochastic frontier efficiency measure on the DEA measure yields (standard error in parentheses), with an $R^2 = 0.395$,

$$RevEff_{sf,n} = \underset{(0.0282)}{-0.0011} + \underset{(0.0606)}{0.4264} * RevEff_{DEA,n} . \tag{I-7}$$

The dependent variable remains the natural logarithm of revenue and the independent variables are time effects and the natural logarithm of employees, oil reserves, and natural gas reserves for each year separately. The hypothesis is that the coefficients on the three inputs are equal across the three time periods. For each set of inputs, a test of this hypothesis can not be rejected at even the 10% level. Thus we proceed with the assumption that the data is poolable.

[21] γ is computed and tested for each of the subsequent stochastic frontier models we estimate, and we find similar results. Table 2 reports γ for each of these models.

Figure I-1: DEA (average) and stochastic frontier revenue efficiency measure



The fact that two techniques using very different approaches yield similar results should increase confidence that the results reflect genuine underlying differences between the firms.[22] In particular, both measures clearly show that the major IOCs are the most revenue efficient, while the NOCs tend to be among the least efficient at raising revenue from inputs of reserves and employees.

[22] An intriguing feature of Figure 1 is that the "outliers" to the regression relationship between the two efficiency measures tend to have a DEA score above their stochastic frontier score. A possible explanation could be that the assumed functional form for the stochastic frontier is inappropriate for these firms.

d. Explaining Inefficiency

Many factors can influence the relative rankings of such a wide variety of firms, especially when the generation of revenue is the only measured objective. For instance, a NOC with non-commercial goals might be expected to be *off* the *revenue* frontier when compared to a firm with no such objectives. In addition, as noted above, different degrees of vertical integration could also affect measured revenue efficiency.

In order to test this hypothesis, we allowed additional variables, such as vertical integration and share of government ownership to influence the frontier using data envelopment analysis. Model 1 in Table I-2 summarizes the average revenue efficiency measures from the basic DEA model presented in the previous section. Model 2 adds vertical integration (defined as petroleum product sales divided by total liquids production) as an input, while Model 3 adds both vertical integration and share of *non-*government ownership.[23] The results show that the measured revenue efficiency of each firm in Model 1 can to a large extent be explained by these included structural and institutional variables.

Table I-2: Summary of firm revenue efficiency, average data envelopment analysis

	All firms	NOC	Major IOC	Others
Model 1	0.394	0.276	0.728	0.449
Model 2	0.622	0.441	0.980	0.726
Model 3	0.768	0.754	0.981	0.754

Including the vertical integration variable in Model 2 corrects for a measurement issue introduced by the manner in which we have defined inputs and outputs, as

[23] Unlike SFA, DEA requires all inputs to have a positive impact on production. Non-government ownership is simply one minus the share of government ownership.

discussed above. A noteworthy result is that accounting for vertical integration moves the five major IOCs toward the frontier. Thus, the data indicates that the corporate structure of a firm is an important feature in the production of revenue.

Adding the non-government ownership share in Model 3 reveals that variable as being responsible for a large amount of the measured revenue inefficiencies that remain in Model 2, particularly for the NOCs. Thus, the DEA analysis confirms the result that government ownership reduces the ability of a firm to produce revenues for a given quantity of inputs.

Although the DEA approach suggests that relative firm efficiency is hindered by government ownership, we employ stochastic frontier approach because it allows us to test the statistical significance of specific characteristics which may impact the ability of a firm to generate revenue. Vertical integration and government ownership share are structural and institutional features that may affect the firm's ability to transform employees and reserves into revenue. Two-tiered pricing also impairs the firm's ability to generate revenue by forcing it to sell some output at subsidized prices.[24] One could view these variables as explaining some of the measured inefficiency, u_n. SFA, therefore, allows us to identify the objectives of the NOC that may not be consistent with a purely profit maximizing firm.

For convenience, we shall refer to the basic stochastic frontier model estimated in the previous section as Model 1sf. Departing from Model 1sf, we systematically add the variables discussed above to determine whether they can explain the estimated deviations from the frontier. For example, Model 2sf adds a measure of vertical integration as an

[24] Other factors may matter in individual cases, but the point of this exercise is to focus on measurable systematic influences on revenue efficiency and more specifically to investigate whether there are any systematic differences between NOC's and other firms.

input. Model 3sf then adds measures of both government share and vertical integration. Following the conclusions of the model developed by Hartley and Medlock (2007), Model 4sf then adds a dummy variable ($2TierP$) for those countries that subsidize domestic prices and an interaction term between the share of government ownership and total employment. Table I-3 presents the estimation results for each of the various stochastic revenue frontier models.[25]

Model 2sf shows that vertical integration is significant in explaining why firms are estimated to be inefficient in generating revenue. The positive coefficient on the vertical integration variable indicates that vertical integration enhances a firm's ability to generate revenue. Prior to controlling for vertical integration, firms with large investments in downstream activities will tend to have a less negative u_n and thus will appear more efficient at generating revenue. As noted above, this is due to the fact that we are not accounting for capital inputs used in refining and marketing operations to produce higher valued products. For example, when we move to Model 2sf, Nippon Oil, which is heavily integrated in downstream activities, actually moves away from the estimated frontier.

The negative coefficient on the government share variable in Model 3sf indicates, as hypothesized at the outset, that government ownership tends to reduce a firm's ability to produce revenue from given inputs. This coefficient summarizes the overall effect of

[25] We also test the hypothesis of poolablility for Model 4sf (see footnote 24). The hypothesis that the coefficient for each input is constant across the three years can only be rejected for vertical integration at the 5% level, but not the 2.5% level. However, this result seems to be driven by increased product sales in 2003 for the two most vertically integrated firms in our sample, Nippon Oil and PTT. Upon dropping these firms, the hypothesis cannot be rejected for any input at the 5% significance level.

government ownership.[26] The remaining model in Table I-3 adds new variables to the set of inputs to help us understand the source of this overall negative effect.

Table I-3: Panel estimation of stochastic frontier

	Model 1sf	Model 2sf	Model 3sf	Model 4sf
ln *L*	0.3786*** 0.0638	0.3761*** 0.0649	0.4300*** 0.0924	0.5600*** 0.0313
ln *OilRsv*	0.1203** 0.0612	0.1795*** 0.0559	0.1801*** 0.0656	0.2438*** 0.0414
ln *NGRsv*	0.1888*** 0.0551	0.1758*** 0.0566	0.2101*** 0.0787	0.1567*** 0.0359
VertInt		0.0903*** 0.0254	0.0868*** 0.0286	0.1071*** 0.0161
GovShare			-0.5254*** 0.1499	2.7658*** 0.6016
2*TierP*				-0.9644*** 0.1260
GovShare * ln *L*				-0.2821*** 0.0590
*year*2003	0.2705*** 0.0247	0.2648*** 0.0253	0.2647*** 0.0257	0.2604*** 0.0262
*year*2004	0.4416*** 0.0250	0.4400*** 0.0255	0.4295*** 0.0261	0.4331*** 0.0263
constant	4.9421*** 0.5960	4.0124*** 0.8202	3.0038** 1.2613	1.3993*** 0.2297
γ	0.9734	0.9678	0.9698	0.9840
95% Conf. Interval of γ	0.9593 to 0.9826	0.9473 to 0.9805	0.9204 to 0.9889	0.9289 to 0.9966
$\chi^2(d)$	519.63	497.40	562.74	2664.44
d	5	6	7	9
Log Likelihood	-78.10	-70.90	-66.14	-41.93
# Observations	237	237	237	237

Model 4sf includes a dummy variable for companies headquartered in countries where domestic prices are subsidized.[27] This enables us to capture the effect of a lower

[26] We also examined the case in which *GovShare* was allowed to differ for importing and exporting firms. The coefficients were not statistically different from each other or the *GovShare* variable in Model 3sf.
[27] Countries for which the 2-tiered pricing dummy was positive are Colombia, Mexico, Russia, China, Thailand, Kazakhstan, Nigeria, Ecuador, Angola, Azerbaijan, Malaysia, Syria, Oman, UAE, Kuwait, Algeria, Indonesia, Saudi Arabia, Iran, and Venezuela.

average sales price for firms likely to be selling a significant amount of output at subsidized domestic oil prices. As discussed above, such subsidies might be imposed to garner political support from a broad constituency. The negative, and highly significant, coefficient on this two-tier pricing variable indicates that subsidized domestic prices do adversely affect the firm's ability to produce revenues.[28]

Government ownership can also lead to excessive employment in a NOC as another way of redistributing resource rents. Model 4sf therefore also includes an interaction term between employment and government share.[29] The estimated coefficient is strongly negative and highly statistically significant. In addition, the three physical variable inputs have a strong and statistically significant positive effect on the production of revenue, while the influence of the vertical integration is not significantly different than Model 3sf.[30]

Model 4sf confirms that government ownership tends to result in a larger workforce than necessary to meet purely commercial objectives. This follows from the coefficients on government share and the interaction term. In particular, if government share is zero, then these variables drop out of the equation. However, the combined effect of government share and the interaction term can be written

---

[28] Although 2003 fuel prices are not available from Metschies, we test the appropriateness of using the 2-tiered pricing dummy to capture the impact of price subsidies using data for 2002 and 2004. The 2-tier pricing variable is replaced with the level of subsidy (if positive). None the coefficients is statistically different from those of Model 4sf. However, the coefficient on *GovShare* becomes insignificant.
[29] We also examined the interaction between *GovShare* and reserves of oil and natural gas reserves. We found these interaction terms to be insignificant. This is consistent with the theoretical model presented in Hartley and Medlock (2007), which predicts ambiguous effects of government ownership on the level of reserves, conditional on the age of the resource.
[30] Model 4sf fits the data reasonably well in terms of its ability to recreate observed firm revenues. However, there are three points, corresponding to the firms SPC and Socar, which are outliers with regard to goodness of fit. These two companies also happen to be the firms for which the data does not cover all three years, while the data that is included appears to be "rounded off". Thus, the fact that the model does not fit these firms very well is likely related to insufficiency of the data set.

$$GovShare*(2.7658-0.2821*\ln L),$$

which is generally negative for firms with a positive government share.[31] The negative coefficient on the interaction term can also be interpreted as implying increased employment has a diminishing positive effect on revenue (that is, has a lower marginal revenue product) the higher is the government share in ownership.[32]

An interesting regularity in Model 4sf is that the Russian firms (regardless of the amount of government ownership) tend to be ranked with low levels of revenue efficiency. This suggests that systematic features of doing business in Russia apart from government ownership negatively affect the ability of Russian firms to generate revenue from a given level of reserves and employment.[33]

Table I-4 summarizes the revenue efficiency scores for the four stochastic frontier models. Adding vertical integration moved firms closer to the frontier, especially for the major international oil companies. The national oil companies, however, moved toward the frontier with the introduction of the government share variable in Model 3sf.

Furthermore, the general tendency is that revenue tends to decrease with an increase in the exercise of government controls. For example, if a firm is forced to sell into a subsidized market, its revenues are impacted negatively. In addition, although an

---

[31] Among all firms with a positive government share in 2004, only two are not subject to two-tier pricing and thus the overall impact of government ownership may be positive. OMV (with 5748 employees and 25% government ownership in 2004) and TPAO (with 5184 employees and 100% government ownership in 2004) have positive coefficients of 0.0810 and 0.3238, respectively. These are statistically positive at the 5% level, but not the 2.5% level.

[32] Following Schmidt and Sickles (1984), we also estimate the random effects stochastic frontier model. The results are consistent with the implications of the four stochastic frontier models using MLE. In particular, the signs are the same and each coefficient is significant at the 1% level in all four models. The fixed effects approach is inappropriate, however, because our analysis contains time-invariant regressors.

[33] To guard against the possibility that the Russian firms are affecting the remaining coefficient estimates, we examined the effect of including a dummy variable for Russian companies in Model 4sf. It had a coefficient of -1.263 and standard error of 0.170. Including this variable did not, however, significantly affect any of the remaining coefficients.

increase in the number of employees tends to increase revenues, firms with full government ownership will generate less revenue for a given level of employment. The largest three firms – PetroChina, Sinopec, and Gazprom – are each fully owned by the government and domestic prices are subsidized. Therefore, the fact that the NOCs move toward the frontier in Model 4sf indicates that these objectives are significant to national oil companies.

Table I-4: Summary of firm revenue efficiency, stochastic frontier analysis

	All firms	NOC	Major IOC	Others
Model 1sf	0.168	0.128	0.600	0.148
Model 2sf	0.253	0.197	0.805	0.233
Model 3sf	0.351	0.329	0.843	0.443
Model 4sf	0.567	0.538	0.893	0.551

e. Concluding remarks

The evidence provided in this paper, using both non-parametric and parametric techniques, supports the notion that non-commercial government objectives can negatively affect the ability of a NOC to earn revenue. In particular, we have uncovered evidence that the government overseers of such firms tend to redistribute resource rents toward domestic consumers of oil products and domestic employees of the firm. Such pressures can alter the investment patterns of the NOC and result in an outcome that can be described as operationally inefficient.

We conclude that the relative revenue inefficiencies of various NOC's, which are observed when one considers only commercial objectives, are largely the result of governments exercising control over the distribution of rents. This is an important finding. If an increasing proportion of global oil and gas resources are under the control of NOC's, it is reasonable to expect that an increasing majority of oil and gas developments will be undertaken with political objectives in mind. This will result in inefficiencies in the production of revenues, and possibly lower levels of production and higher prices, than would occur under commercial development.

f. References

Afriat, S.N. (1972). "Efficiency Estimation of Production Functions," *International Economic Review*, 13: 568-598.

Aigner, D.J. and S.F. Chu (1968). "On Estimating the Industry Production Function," *American Economic Review*, 58(4): 826-839.

Aigner, D.J., C.A.K. Lovell, and P. Schmidt (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics*, 6(1): 21-37.

Alchian, A.A. (1965). "Some Economics of Property Rights," *Il Politico*, 30(4): 816-829.

Al-Obaidan, A.M. and G.W. Scully (1991). "Efficiency Differences between Private and State-owned Enterprises in the International Petroleum Industry," *Applied Economics*, 23: 237-246.

Banker, R.D., A. Charnes, and W.W. Cooper (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science*, 30(9): 1078-1092.

Banker, R.D. (1993). "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation," *Management Science*, 39(10): 1256-1273.

Battese, G.E., and T.J. Coelli (1988). "Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data," *Journal of Econometrics*, 38(3): 387-399.

Battese, G.E., and T.J. Coelli (1992). "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farms in India," *Journal of Productivity Analysis*, 3(1-2): 153-169.

Boardman, A.E. and A.R. Vining (1989). "Ownership and Performance in Competitive Environments: A Comparison of the Performance of Private, Mixed, and State-Owned Enterprises," *Journal of Law and Economics*, 32(1): 1-33.

Boles, J.N. (1966). "Efficiency Squared – Efficient Computation of Efficiency Indexes," *Proceedings of the 39th Annual Meeting of the Western Farm Economic Association*, 137-142.

Charnes, A., W.W. Cooper and E. Rhodes (1978). "Measuring the Efficiency of Decision Making Units," *European Journal of Operations Research*, 2: 429-444.

Coelli, T.J. (1996). "A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program", CEPA Working Paper 96/8, Department of Econometrics, University of New England, Armidale NSW Australia.

Coelli, T., D.S.P. Rao, and G.E. Battese (1999). *An Introduction to Efficiciency and Productive Analysis*. Boston: Kluw3er Academic Publishers.

Cooper, W.W. and K. Tone (1997). "Measures of Inefficiency in Data Envelopment Analysis and Stochastic Frontier Estimation," *European Journal of Operational Research*, 99(1): 72-88.

Cornwell, C., P. Schmidt, and R.C. Sickles (1990). "Production Froniters with Cross-Sectional and Time-Series Variation in Efficiency Levels," *Journal of Econometrics*, 46(1-2): 185-200.

Côte, D. (1984). "Firm Efficiency and Ownership Structure: The Case of the U.S. Electric Utilities Using Panel Data," *Annals of Public and Cooperative Economics*, 60: 431-450.

Debreu, G. (1951). "The Coefficient of Resource Utilization," *Econometrica*, 19(3): 273-292.

Green, W.H. (2003). *Econometric Analysis*. 5th ed. New Jersey: Prentice-Hall.

Energy Intelligence (2004, 2005, 2006) *The Energy Intelligence Top 100: Ranking the World's Oil Companies*.

Energy Information Administration (2004, 2006). *International Energy Annual*, "Table 2.2 World Crude Oil Production, 1980-2004".

Energy Information Administration (2007). "World Proved Crude Oil Reserves, January 1, 1980- January 1, 2007 Estimates".

Farrell, M.J. (1957). "The Measurement of Productive Efficiency", *Journal of Royal Statistical Society*, 120(3): 11-48.

Gong, B.H. and R.C. Sickles (1992). "Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data," *Journal of Econometrics*, 51(1-2): 259-284).

Harris, M. and A. Raviv (1978). "Some Results on Incentive Contracts with Applications to Education and Employment, Health Insurance, and Law Enforcement," *American Economic Review*, 68(1): 20–30.

Hartley, P.R. and K.B. Medlock III (2007). "A Model of the Operation and Development of a National Oil Company," James A, Baker III Institute for Public Policy working paper.

Jensen, C.M. and W.H. Meckling (1979). "Rights and Production Functions: An Application to Labor-Managed Firms and Codetermination," 52(4): 469–506.

Jondrow, J., C.A.K. Lovell, I.S. Materov, and P. Schmidt (1982). "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics*, 9(2-3): 233-238.

Kumbhakar, S.C. (1987). "The Specification of Technical and Allocative Inefficiency in Stochastic Production and Profit Frontiers," *Journal of Econometrics*, 34(3): 335-348.

Kumbhakar, S.C. and C.A.K. Lovell (2000). *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Laffont, J.J. and J. Tirole (1991). "Privatization and Incentives," *Journal of Law, Economics and Organization*, 7(Special issue): 84–105.

Koopmans, T.C. (1951). "An Analysis of Production as an Efficient Combination of Activities." In Koopmans (ed.), *Activity Analysis of Production and Allocation*. Cowles Commission for Research in Economics, Monograph 13. New York: John Wiley and Sons, Inc.

Metschies, G.P. (2003). *International Fuel Prices 2003 - 3rd Edition.* Available at www.internationalfuelprices.com.

Metschies, G.P. (2005). *International Fuel Prices 2005 – 4th Edition.* Available at www.internationalfuelprices.com.

Meeusen, W. and J. van dan Broeck (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error," *International Econometric Review*, 18(2): 435-444.

Parker, D. (1995). "Privatization and Agency Status, Identifying the Critical Factors for Performance Improvement," *British Journal of Management*, 6: 29-43.

Petroleum Intelligence Weekly (2006), "PIW's Top 50: How the Firms Stack Up," *Petroleum Intelligence Weekly: Special Supplement*, 45(51): 2-3.

Pitt, M. and L.F. Lee (1981). "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics*, 9:715-723.

Ruggiero, J. (2007). "A Comparison of DEA and Stochastic Frontier Model Using Panel Data," *International Transactions in Operational Research*, 14(3): 259-266.

Schmidt, K.M. (1996). "The Cost and Benefits of Privatization: an Incomplete Contracts Approach," *Journal of Law, Economics and Organization*, 12(1): 1-24.

Schmidt, P. and R.C. Sickles (1984). "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics*, 2(4): 367-374.

Shephard, R.W. (1953). *Cost and Production Functions*. Princeton: Princeton University Press.

StataCorp (2007). *Stata Statistical Software: Release 9*. College Station, TX: StataCorp LP.

Vickers, J. and G. Yarrow (1991). "Economic Perspectives on Privatization," *The Journal of Economic Perspectives*, 5(2): 111-132.

Villalonga, B. (2000). "Privatization and Efficiency: Differentiating Ownership Effects from Political, Organizational, and Dynamic Effects," *Journal of Economic Behavior & Organization*, 42: 43-74.

II. Essay 2 – Energy Demand and Economic Growth: Relationships and Implications

a. Introduction

Economic development causes aggregate production and consumer demand to evolve over time. Energy demand, which is derived from the needs of producers and consumers to use energy as an input, will exhibit a pattern of growth that mirrors structural changes induced by economic development. In general, energy intensity, defined as energy use per unit output, appears to increase during early stages of development. However, it is often claimed that processes of *dematerialization* (a shift in production toward goods and services that require less energy and material inputs) and *saturation* (a bound on the utilization rate of energy using capital) cause energy intensity to eventually decrease, resulting in a non-monotonic relationship between economic development and energy demand. We investigate these claims using a panel of 50 countries spanning all levels of development.

One of the innovations in our study is that we examine the evolution of energy consumption by energy source and end-use sector. The unique combinations of energy commodity and sector are labeled commodity-sectors throughout this paper. We propose that an additional factor, one we call *capital utilization*, causes short-term increase in the utilization of inefficient capital and therefore energy intensity in periods when economic growth exceeds the long-run average. Also, a decrease in the utilization of inefficient capital in times of economic contraction causes a short-term decrease in energy intensity. To accommodate this factor, we allow energy demand to be positively correlated with economic growth in the short-term during all stages of economic development.

Conversely, in the long-run, we allow dematerialization and saturation to cause energy use to be negatively correlated with economic growth in developed economies.

Our model is used to compare the relationship between economic development and energy demand by commodity and sector. Our results by sector are consistent with the impact of dematerialization and saturation. However, we conclude that electricity, as compared to other direct-use fuels, exhibits the strongest evidence of a non-monotonic relationship between economic development and energy demand. Moreover, we find the long run income elasticity of electricity demand to be higher than direct-use fuel in earlier stages of economic development.

Our paper proceeds as follows. Section b discusses the structural composition of energy demand. This is followed by a discussion of empirical research on the topic in section c. Our model is developed and presented in section d. Section e discusses the data we use in our study. The following three sections, f through h, discuss and test the specification of our model with respect to the set of endogenous regressors, the nature of the time and country-specific effects and heterogeneity of the error covariance matrix. Section i presents the parameter estimates and as well as simulated demand paths for a hypothetical country for each energy commodity-sector. Sections j and k then analyze and compare demand by sector and by commodity, respectively. The paper concludes in section k.

b. Structural Components of Energy Demand

The composition of production has long been a topic of interest in the field of development economics. Kuznets (1971) and Chenery and Syrquin (1975) first described the pattern of economic development in relation to the relative shares of aggregate output by sector. In the first stage of economic development, industrial production begins to grow and displace agriculture as the single largest share of output. As consumer income grows, however, consumer demands shift and growth of the service sector dominates that of all other sectors of aggregate output.

Since energy is a derived demand and energy requirements vary by sector, aggregate energy consumption can evolve in complicated ways with economic development. In particular, a predominately agricultural economy with largely subsistence farming neither requires large quantities of energy nor has significant levels of disposable income. Thus, energy consumption is initially low. However, as the nation transitions first into more market-based agriculture with more trade in agricultural products and later into an industrialized economy, infrastructure such as roads, ports, and buildings must be developed. The production of inputs (such as steel and cement) required for the development of necessary infrastructure is intensive in the use of energy and other raw materials. As a result, energy intensity increases rapidly during industrialization. Then as disposable income increases, demand for consumer durables and manufactured goods also increases. Industrial production transitions from heavy to light industry to accommodate increasing consumer demand for manufactured goods. Furthermore, services, which are relatively less energy intensive than industrial production, grow relative to total output. This process, known as the dematerialization of production, decreases the intensity of use for energy and other raw materials.[34]

The saturation effect, on the other hand, relates specifically to the changing pattern of consumer activity as personal wealth increases. Because many of the manufactured goods that consumers buy (air conditioners, heaters, refrigerators and

[34] See Bernardini and Galli (1993) for more details on the process of dematerialization.

personal vehicles, for example) consume energy and require expanded retail and transportation services, the share of total energy consumption in the residential, transportation and commercial sectors begins to grow faster than industrial demand for energy. However, there is a limit, or level of saturation, at which per capita energy demand from these sectors can no longer increase with economic growth. For example, while the demand for personal vehicles increases with income, an individual may not operate a vehicle more than 24 hours a day. Thus the demand for transportation fuels will likely have some upper bound relative to individual consumption. This saturation effect is also likely to be present in the residential and commercial sectors due to the fact that the utilization rate of capital is bounded above by 100%. It should be noted, however, that energy consumption at the national level may continue to grow as population increases.

Additionally, we propose a third factor that causes short-term increases in energy intensity in periods of rapid growth and decreases in energy intensity in periods of economic contraction. We label this process capital utilization as it relates to economic growth and the use of less efficient capital to increase production.

Consumption of energy in the industrial sector is derived primarily from using energy as an input into the production process. Capital stock, however, is generally fixed in the short-term, and it is difficult to change the type of energy used to power a given piece of equipment. Assuming industrial producers are efficient with respect to cost minimization, it is likely that producers with the more energy efficient capital stock (or relatively low use of energy per unit output) will be lower on the supply curve. Thus, any increase in output must be met by producers owning less energy efficient capital.

Holding all else equal, energy intensity (defined as energy use per unit output) will increase as output expands. Similarly, inefficient capital utilization is likely to decrease as output decreases, causing energy intensity to also decrease. Energy demand is, therefore, likely to be positively correlated in the short-term with economic growth. However, capital stock is not fixed in the long run, and in efficient markets, additions will begin to reduce the share of inefficient capital stock and erode the short run increase in energy intensity.

c. Analyzing Energy Demand

Total energy demand evolves to reflect changes in the amount and composition of aggregate production and consumer demand. A model of energy consumption, therefore, must be flexible with respect to economic development. In particular, the elasticity of energy demand must not be constant at all levels of income, rather it must be allowed to decrease with economic growth.

In modeling energy demand as a function of economic development, it is necessary to model each sector separately because energy demand is derived from needs, which vary by sector. In general, the consumption of energy in the industrial sector is derived from the production process, whereas energy demands in the residential and commercial sectors are derived from the powering of equipment and devices, and the heating and cooling of one's home or office. In contrast, the consumption of energy (primarily gasoline and diesel fuel) in the transportation sector is derived from the quantity and length of time that people and goods are transported. It is hypothesized that each sector is characterized by a decreasing elasticity of demand with respect to income. However, dematerialization is most likely to occur in the industrial sector, where as saturation is more likely to occur in the transportation and residential and commercial sectors. Furthermore, short-term increases in energy intensity as a result of capital utilization are likely characteristic of the transportation and industrial sectors. If we assume both labor and energy are variable inputs, then less efficient capital stock and additional employees will be used to increase output, resulting in more transportation of the workforce and finished goods.

There are three notable studies (Galli (1998), Judson, Schmalensee, and Stoker (1999) and Medlock and Soligo (2001)) which use panel data and allow income elasticity of total energy consumption per capita to decline with economic growth. Judson, Schmalensee, and Stoker (1999) use panel data for 123 countries covering developed and less developed economies for the years 1950 through 1991. The authors estimate per capita energy consumption in five sectors as a function of per capita income with time and country-effects. The five sectors are defined as industry and construction, transportation, household and other, the energy sector, and non-energy uses.

Since energy price data are not available for many countries, time effects are used to capture the influence of energy prices. Although the authors find a general pattern of declining income elasticity, the exclusion of energy prices (due to a lack of data) may introduce bias due to omitted variables. Including both prices and income allows the impact of dematerialization and saturation to be distinguishable from the impact of prices. Otherwise, a positive price shock resulting in reduced energy demand could be mistaken as an income effect due to dematerialization and saturation. Furthermore, the inclusion of time effects may capture short-term increases in energy intensity caused by higher utilization of inefficient capital during periods of rapid economic growth. This would also bias the model and lead to lower estimates of short run income elasticity.

Galli (1998) estimates trends in long-term energy intensity using an error correction model for a panel of 10 developing Asian economies from 1973 to 1991. A quadratic function in log income is used and the author finds the long run coefficient on squared income to be negative and significant. Although Galli's study is smaller in scope than the study by Judson, Schmalensee, Stoker (1999), the estimated coefficient on price was found to be significant.[35]

The study by Medlock and Soligo (2001) is described by the authors as a hybrid of Judson, Schmalensee, and Stoker (1999) and Galli (1998). For a panel of 28 countries from 1978 to 1995, the authors model (in logs) the per capita long run equilibrium demand for energy by sector as a function of real energy prices and real output per capita. The possibility of a non-constant income elasticity of demand is captured by including a log-quadratic variable for income per capita. The authors incorporate short run dynamics by employing the Koyck (1954) partial adjustment mechanism and find evidence that income elasticity declines with economic development. In particular, energy demand per capita is maximized at income levels near $16,000 in the industrial and other sector, $34,000 in the residential and commercial sector, and $533,000 in the transportation sector, where real per capita income is measured as GDP per capita denominated in 1995 dollars adjusted for purchasing power parity.

We expand upon the research of Medlock and Soligo (2001) in several ways. First, we incorporate an additional parameter to capture the impact of capital utilization

[35] With respect to capital utilization, an error correction model would suffer less bias than a model in levels with time effects. However, the marginal impact on changes in energy consumption with respect to a change in output (as measured by GDP) would likely be non-linear .

which causes short-term increases in energy intensity in periods of economic growth. Second, energy demand by sector is modeled separately for electricity and all other energy sources, the latter of which we label as direct-use fuel. There is evidence to suggest that income elasticity of electricity demand also decreases with income, but not necessarily in the same way as direct-use fuel (see Hankinson and Rhys (1983), Ang (1988) and Nilsson (1993), for example). Ferguson, Wilkinson and Hill (2000) show that "wealthy countries have a stronger correlation between electricity use and wealth creation than do poor countries and that, for the global economy as a whole, there is a stronger correlation between electricity use and wealth creation than there is between total energy use and wealth."

While large electric generators that convert direct-use fuel into electricity benefit from economies of scale, the use of electricity in the industrial, residential, and commercial sectors requires significant investment in a large, integrated transmission and distribution system. Moreover, electricity and direct-use fuel are not perfect substitutes as a result of differing physical characteristics. In particular, equipment is generally designed to use only one type of fuel. Even in the longer run, the energy sources are not perfectly substitutable because of differing cost characteristics. Transportation of electricity is relatively costly due to high energy loss and more especially because it requires point-to-point connections via transmission lines. In contrast, shipping costs for direct-use fuels such as petroleum products (shipped by road) and coal (shipped by rail) are relatively low and while the pipeline networks for shipping natural gas also are expensive, transmission losses are substantially lower than for electricity. Finally, while technologies are being developed to improve electricity storage, pumped storage is currently the only effective method and is primarily used to smooth intraday price fluctuations. Storage of direct-use fuel however is relatively cheap and widely available and can arbitrage prices over longer periods of time. In summary, electricity markets are infrastructure intensive in that they require a vast transmission and distribution network. Additionally, the necessity to meet load without the use of storage requires a significant amount of excess capacity that remains idle most of the time. There is, therefore, a significant trade-off between electricity and direct-use fuel in the industrial, residential and commercial sectors.

Third, our study differs from Medlock and Soligo (2001) in that we do not model energy consumption per capita as a function of per capita output. Instead we have chosen to model total energy consumption as a function of aggregate output and population. This allows us to separate the elasticity of energy demand with respect to income and population.

Finally, the data set used by Medlock and Soligo (2001) contains statistics for 28 countries from 1978 to 1995. Although similar, the dataset created for our study is unique in its inclusion of energy price statistics by energy source. We have collected data for 50 countries for the years 1990 to 2004. The majority of the additions are low and middle income countries, many from Latin America. It should be noted that these 50 countries account for 83.9% of all commercial energy consumption in the world from 1990 to 2004. Table II-1 lists our sample countries and notes those not included in Medlock and Soligo (2001).[36] Income data, described below, are included to highlight the number of low and middle income countries added.

[36] Vietnam was included in Medlock and Soligo (2001). We chose to eliminate Vietnam from our sample due to a lack of sufficient price data.

Table II-1: Sample countries and 2004 income

Country	GDP per capita *2000 PPP $*	Country	GDP per capita *2000 PPP $*	Country	GDP per capita *2000 PPP $*
Pakistan	2,004	Uruguay*	8,374	Germany*	25,945
Bolivia*	2,456	Malaysia	9,374	France	26,872
India	2,851	Mexico	9,385	Japan	27,114
Nicaragua*	3,208	South Africa*	9,533	Australia	27,840
Indonesia	3,282	Chile*	10,168	Finland	28,116
Ecuador*	3,740	Argentina*	11,750	Sweden	28,276
Egypt*	3,747	Poland*	11,913	Belgium	28,379
Sri Lanka*	3,914	Slovak Rep.*	13,329	Netherlands	28,819
Paraguay*	4,106	Hungary*	15,254	Canada	29,136
Philippines*	4,431	Czech Rep.*	17,270	United Kingdom	29,231
Peru*	5,122	Portugal	18,172	Denmark	29,409
China	5,490	Korea	18,934	Austria	29,662
Colombia*	6,275	Greece	20,143	Switzerland*	31,310
Panama*	6,473	New Zealand*	22,025	Ireland	33,194
Turkey	7,055	Israel*	22,265	Norway	36,282
Brazil	7,406	Spain	23,757	United States	36,451
Thailand	7,453	Italy	25,579		

* denotes additions to Medlock and Soligo (2001)

Figure II-1 to Figure II-8 plot (in logs) energy consumption as a function of total GDP and population for total final energy consumption as well as the five commodity-sectors we model. The three end-use sectors are defined as transportation, industrial, and residential, commercial and agriculture (or RCA for short). The two energy commodities are electricity and direct-use fuels. The unique combinations of energy sources and end-use sectors will be frequently referred to as commodity-sectors below.[37] These plots demonstrate that our sample spans all levels economic development and population throughout the 15 years of our model.

[37]We test and reject the hypothesis that the model of energy consumption can be pooled by energy commodity and sector into one model of total energy consumption. This is presented in section i.

Figure II-1: GDP and total energy demand

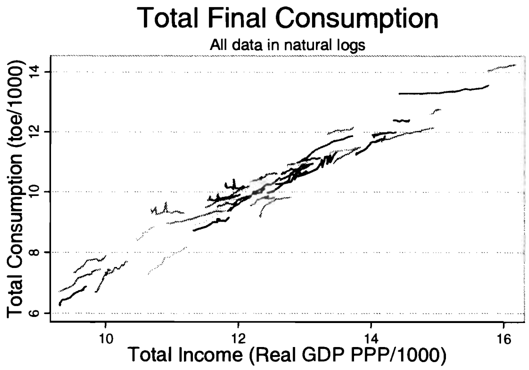


Figure II-2: GDP and industrial electricity demand

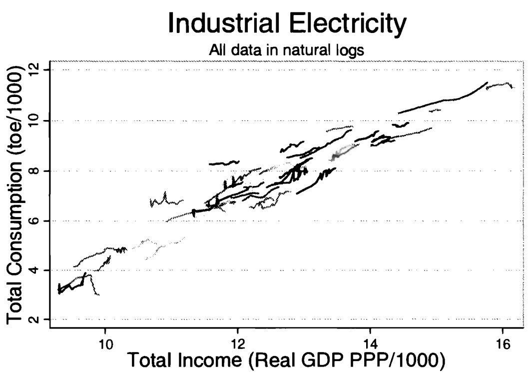


Figure II-3: GDP and industrial direct-use energy demand

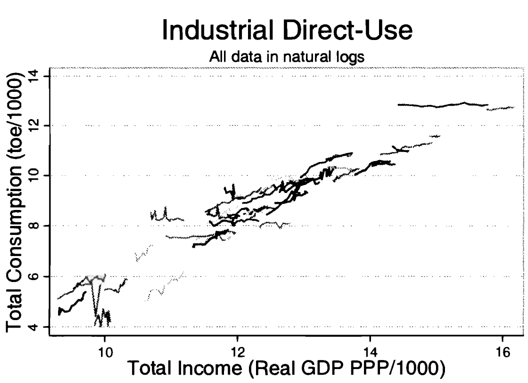


Figure II-4: GDP and transportation fuel demand

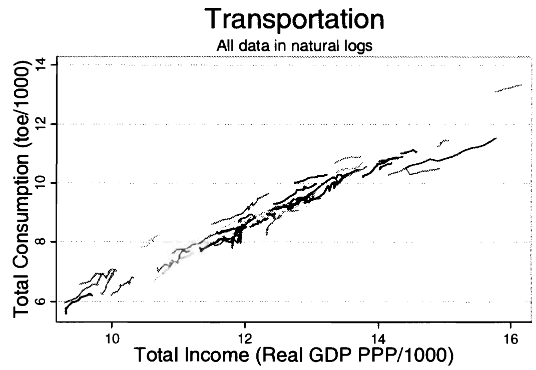


Figure II-5: GDP and RCA electricity demand

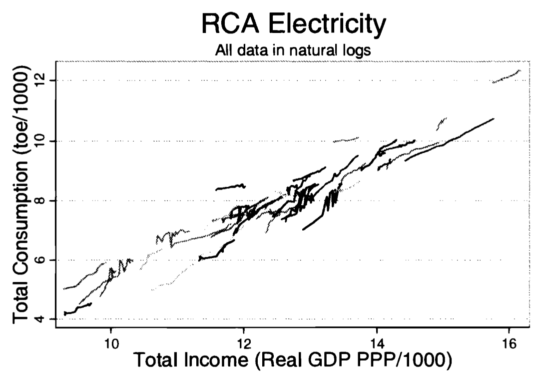


Figure II-6: GDP and RCA direct-use energy demand

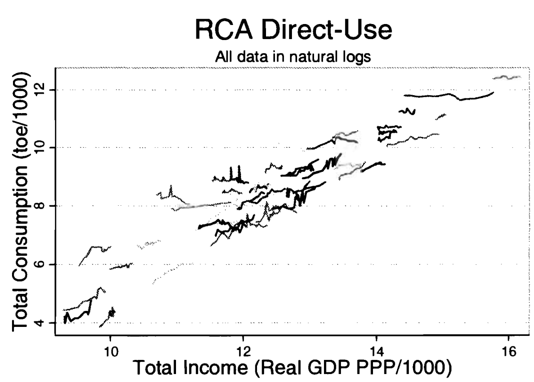


Figure II-7: Population and total energy demand

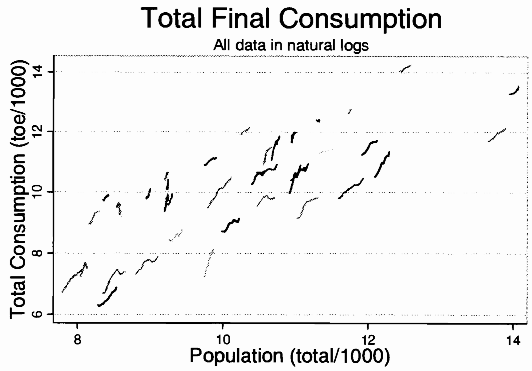


Figure II-10: Population and transportation fuel demand

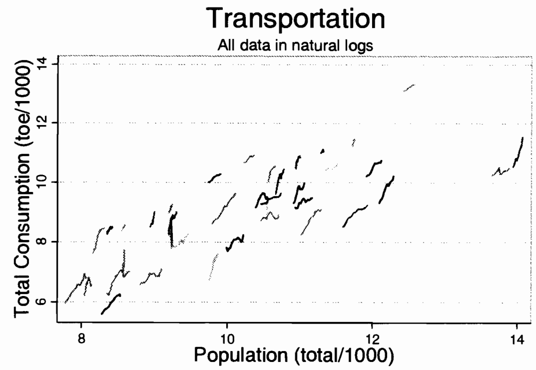


Figure II-8: Population and industrial electricity demand

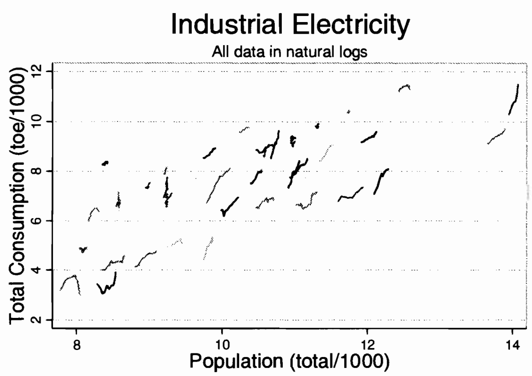


Figure II-11: Population and RCA electricity demand

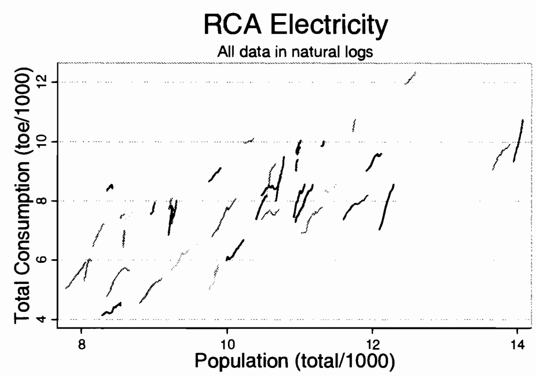


Figure II-9: Population and industrial direct-use energy demand

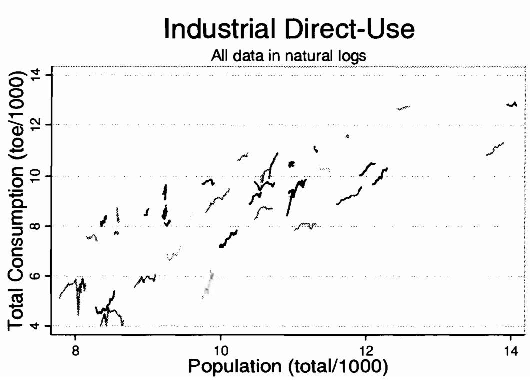
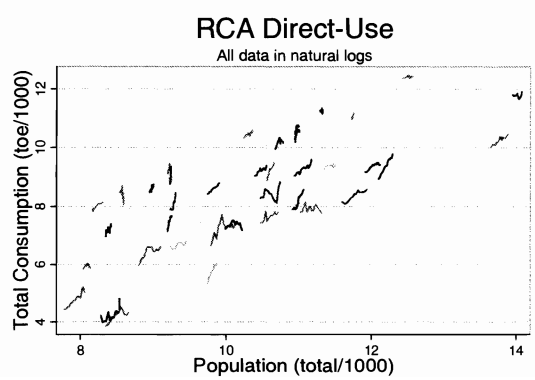


Figure II-12: Population and RCA direct-use energy demand



d. Commodity-Sector Model of Energy Consumption

Letting i and t, respectively, denote country and time, consider a basic model of energy consumption where long run equilibrium demand ($E^*_{i,t}$) is a function of energy price ($p_{i,t}$), population ($pop_{i,t}$), aggregate output ($Y_{i,t}$) and technology ($\tau_{i,t}$). This can be written as

$$E^*_{i,t} = f(p_{i,t}, pop_{i,t}, Y_{i,t}, \tau_{i,t}) . \tag{II-1}$$

For the purpose of this study, we disaggregate total energy consumption by end-use sector (j) and energy commodity (k) such that

$$E^*_{i,t} = \sum_j \sum_k E^*_{i,j,k,t} . \tag{II-2}$$

As in Judson, Schmalensee and Stoker (1999), as well as Medlock and Soligo (2001), we include aggregate output in our model rather than sector and commodity-specific output, which would have been denoted $Y_{i,j,k,t}$. The reason is that GDP data is not available at this level of disaggregation.

Assume every country has access to the same technology, and installed technology is a function of only energy prices, the level of domestic output and population. Thus $\tau_{i,j,k,t} = \tau(p_{i,j,k,t}, pop_{i,t}, Y_{i,t})$ and

$$\begin{aligned} E^*_{i,j,k,t} &= f(p_{i,j,k,t}, pop_{i,t}, Y_{i,t}, \tau(p_{i,j,k,t}, pop_{i,t}, Y_{i,t})) \\ &= \hat{f}(p_{i,j,k,t}, pop_{i,t}, Y_{i,t}) \end{aligned} . \tag{II-3}$$

Therefore, one need not model technology explicitly. While this assumption is made to simplify the model, we later test its validity by including time effects (in the form of dummy variables for each year) in the estimation of the model. An inability to reject the hypothesis that the year dummies are jointly equal to zero leads us to conclude our results are not biased by an omission of modeling technology explicitly.

Also, a large share of energy consumption is derived from the demand for space-heating and cooling. As a result, weather (specifically temperature) has a significant impact on total energy demand. To incorporate temperature into our model, we use total heating degree days squared per year for each country.[38, 39] To define heating degree days, let d be a day in year $t \in \{1, 2, ..., T\}$ and $temp_d$ be the average of the maximum and minimum temperatures, measured in degrees Fahrenheit. The number of heating degree days on day d in country i is defined as $HDD_{i,d} = \max\{65 - temp_{i,d}, 0\}$. Thus, including heating degree days in a model of energy demand would allow one to estimate the elasticity of demand with respect to one degree if the temperature is below 65 degrees.[40] However, we find elasticity of demand is nonlinear with respect to temperature and include a parameter of heating degree days squared. Specifically, we include

$$hdd_{i,t}^2 = \left(\frac{\sum_{d \in t} HDD_{i,d}}{1000} \right)^2 \qquad \text{(II-4)}$$

[38] In a fixed effect setting for panel data, omitting heating degree days increases the individual effects to capture that country's average demand for heating. However, in this specification, the clustering of countries across continents renders the time effects in the RCA sector significant because of year-specific heterogeneity resulting from regional winter weather patterns. Explicitly including heating degree days does not eliminate this problem. However, including the squared term causes the joint time effects to be insignificantly different than zero.

[39] We tested and rejected the significance of cooling degree days (in levels and squared) in the model. In the fixed effect setting, omitting cooling degree days increases the individual intercept for countries with significant cooling demand. Where as the exclusion of heating degree day data caused time effects to capture regional cold weather patterns, the exclusion of cooling degree data did not have the same result for hot weather patterns.

[40] The use of heating degree days is fairly common within the energy industry. Similarly, cooling degree days and total degree days are often used. However, we found cooling degree days to be insignificant in each commodity-sector of our model.

to capture marginally increasing demand for energy for every degree of colder temperature.

To allow income elasticity of demand to be non-constant with respect to GDP, we adopt an energy demand function similar to the one specified by Medlock and Soligo (2001).[41] The function we use is

$$E^*_{i,j,k,t} = A_{i,j,k,t} p_{i,j,k,t}^{b_{1,j,k}} pop_{i,t}^{b_{2,j,k}} Y_{i,t}^{b_{3,j,k}} \exp(y_{i,t})^{b_{4,j,k}} \exp(hdd^2_{i,t})^{b_{5,j,k}} \tag{II-5}$$

where $y_{i,t} = (Y_{i,t}/1000)/pop_{i,t}$ is simply output per capita (divided by 1000) and the parameters for each sector and energy source are unique. Taking the natural logarithm of (II-5) yields

$$\begin{aligned} \ln E^*_{i,j,k,t} &= \ln A_{i,j,k,t} + b_{1,j,k} \ln p_{i,j,k,t} + b_{2,j,k} \ln pop_{i,t} \\ &+ b_{3,j,k} \ln Y_{i,t} + b_{4,j,k} y_{i,t} + b_{5,j,k} hdd^2_{i,t} \end{aligned}. \tag{II-6}$$

Time effects and country-specific effects are embodied in the variable $A_{i,j,k,t}$. Denoting the individual country effects for each energy commodity-sector as $a_{i,j,k}$ and the time effects for each sector and energy source as $\theta_{j,k,t}$, then

$$\begin{aligned} \ln E^*_{i,j,k,t} &= a_{i,j,k} + \theta_{j,k,t} + b_{1,j,k} \ln p_{i,j,k,t} + b_{2,j,k} \ln pop_{i,t} \\ &+ b_{3,j,k} \ln Y_{i,t} + b_{4,j,k} y_{i,t} + b_{5,j,k} hdd^2_{i,t} \end{aligned} \tag{II-7}$$

is the long-run equilibrium demand function.

Note that several countries in our sample are in warm climates where the average temperature does not drop below 65. Because the natural logarithm of zero is undefined, we use heating degree days squared ($hdd^2_{i,t}$) in levels in equation (II-7). Also, per capita

---

[41] Medlock and Soligo (2001) use the function $ec^*_{i,j,t} = A_{i,j} p_{i,j,t}^{b_{1,j}} y_{i,t}^{b_{2,j} + b_{3,j} \ln y_{i,t}}$ where $ec^*_{i,j,t}$ denotes energy consumption per capita in country i, in sector j, at time t. $A_{i,j}$, $p_{i,j,t}$ and $y_{i,t}$ are the country and sector specific intercept, price and income per capita, respectively.

income is included in levels to avoid the problem of multicollinearity with total income and population.

Given the demand function specified in equation (II-7), the long-run income elasticity of energy demand is

$$\begin{aligned}\varepsilon^{*Y}_{i,j,k,t} &= \frac{\partial \ln E^{*}_{i,j,k,t}}{\partial Y_{i,t}} \times Y_{i,t} \\ &= b_{3,j,k} + b_{4,j,k}\, y_{i,t} \\ &= b_{3,j,k} + \frac{b_{4,j,k} Y_{i,t}}{1000 * pop_{i,t}}\end{aligned} \tag{II-8}$$

for each, country, sector and commodity. If $b_{3,j,k} > 0$ and $b_{4,j,k} < 0$, income elasticity is initially positive and decreases as per capita income grows. Moreover, the relationship between demand and economic is non-monotonic and the demand maximizing level of income can be solved by setting (II-8) to zero and solving for $y_{i,t}$.

$$y^{*\max} = \frac{-b_{3,j,k}}{b_{4,j,k}} \tag{II-9}$$

Furthermore, the long run elasticity with respect to population is

$$\begin{aligned}\varepsilon^{*pop}_{i,j,k,t} &= \frac{\partial \ln E^{*}_{i,j,k,t}}{\partial pop_{i,t}} \times pop_{i,t} \\ &= b_{2,j,k} - b_{5,j,k}\, y_{i,t} \\ &= b_{2,j,k} - \frac{b_{5,j,k} Y_{i,t}}{1000 * pop_{i,t}}\end{aligned} \tag{II-10}$$

for each sector and energy source. If $b_{2,j,k} > 0$ and $b_{5,j,k} < 0$, elasticity with respect to population is increasing with economic growth.

Finally, the model must incorporate short run dynamics. We opt to implement the Koyck (1954) partial adjustment mechanism,

$$\left(\frac{E_{i,j,k,t}}{E_{i,j,k,t-1}}\right)=\left(\frac{E^{*}_{i,j,k,t}}{E_{i,j,k,t-1}}\right)^{\gamma_{j,k}}, \tag{II-11}$$

which describes the adjustment to the long-run equilibrium for each sector and energy source.[42] Taking the natural logarithm of this expression and rearranging the terms yields

$$\ln E_{i,j,k,t}=\gamma_{j,k}\ln E^{*}_{i,j,k,t}+\left(1-\gamma_{j,k}\right)\ln E_{i,j,k,t-1}. \tag{II-12}$$

In this specification, contemporaneous demand is a function of the long run equilibrium (i.e. contemporaneous price, population, total GDP, GDP per capita and heating degree days squared) and lagged energy demand (which represents the stock of energy using capital). The parameter $\gamma_{j,k}$, which is bounded by zero and one, represents the portion of demand that responds to contemporaneous factors. If $\gamma_{j,k}=1$, capital stock and demand adjust instantly to the long run equilibrium. However, if $\gamma_{j,k}=0$, then capital stock is permanently fixed and the adjustment is infinitely long.

This flow adjustment model, however, does not capture the change in the utilization rate of capital and energy intensity in times of rapid economic expansion or contraction if $\gamma_{j,k}<1$. Recall that consumption of energy in the industrial sector is derived primarily from using energy as an input into the production of goods. If output (as measured by total GDP in our model) grows at a constant rate, g, then additions to capital stock can converge in the long run to some constant rate, k, such that the short run change in capital utilization and energy intensity *equal* zero in every period. Thus we assume $k=f\left(g\right)$ in this case. Alternatively, if the output growth rate is random, g_t,

[42] Baltagi and Griffin (1997) also use a similar adjustment mechanism in their dynamic model of gasoline demand. Galli (1998), however, uses an error correction model. We choose to use the Koyck (1954) partial adjustment mechanism to be consistent with Medlock and Soligo (2001).

and capital additions are fixed at k in the short run, then capital utilization and energy intensity must vary in the short-run. However, if g_t averages g in the long run, then additions to capital stock at constant rate of k cause changes in capital utilization and energy intensity to *average* zero in the long run. In this case, there are periods when the change in output exceeds the long run average growth rate , $g_t > g$, and capital additions, $k < f\left(g_t\right)$, are insufficient to prevent an increase in energy intensity and capital utilization in the short run as less efficient capital is used to increase production. When output growth falls below the long run average, $g_t < g$ and $k > f\left(g_t\right)$, utilization of inefficient capital decreases and energy intensity decreases.

In our paper, we assume that capital additions are a function of average total GDP growth for each country, g_i and are fixed at a constant rate, k_i . Letting g_i equal the long run average growth rate for each country, then $k_i = f\left(g_i\right)$ for each country i. Furthermore, we assume the change in capital utilization is a function of total GDP growth and capital additions, $K_{i,t} = f\left(g_{i,t} - k_i\right) = f\left(g_{i,t} - g_i\right)$. Thus, the change in capital utilization is a function of the deviation from the long run average growth rate. It is computed by taking the difference between the change in log GDP in period t and the average change in log GDP in the previous 10 years. Formally,

$$\begin{aligned} K_{i,t} &= \left(\ln Y_{i,t} - \ln Y_{i,t-1}\right) - \frac{1}{10}\sum_{s=t-10}^{t-1}\left(\ln Y_{i,s} - \ln Y_{i,s-1}\right) \\ &= \left(\ln Y_{i,t} - \ln Y_{i,t-1}\right) - \frac{1}{10}\left(\ln Y_{i,t-1} - \ln Y_{i,t-11}\right) \end{aligned} \quad . \qquad \text{(II-13)}$$

where over a sufficient length of time, the average deviation from the long run GDP growth rate will converge to zero.[43]

Adding $K_{i,t}$ to (II-12), our short run model becomes

$$\ln E_{i,j,k,t} = \gamma_{j,k} \ln E^{*}_{i,j,k,t-1} + \delta_{j,k} K_{i,t} + \left(1-\gamma_{j,k}\right) \ln E_{i,j,k,t-1}. \qquad \text{(II-14)}$$

Substituting the long-run equilibrium demand function described by equation (II-7) into (II-14) yields

$$\begin{aligned} \ln E_{i,j,k,t} &= \gamma_{j,k} a_{i,j,k} + \gamma_{j,k}\theta_{j,k,t} + \gamma_{j,k} b_{1,j,k} \ln p_{i,j,k,t} + \gamma_{j,k} b_{2,j,k} \ln pop_{i,t} + \\ &\gamma_{j,k} b_{3,j,k} \ln Y_{i,t} + \gamma_{j,k} b_{4,j,k} y_{i,t} + \gamma_{j,k} b_{5,j,k} hdd^{2}_{i,t} + \delta_{j,k} K_{i,t} + \left(1-\gamma_{j,k}\right) \ln E_{i,j,k,t-1} \end{aligned}, \qquad \text{(II-15)}$$

the short run dynamic model of energy consumption for each country, sector and energy source. Let $\gamma_{j,k} a_{i,j,k} = \alpha_{i,j,k}$, $\gamma_{j,k}\theta_{j,k,t} = \phi_{j,k,t}$, and $\gamma_{j,k} b_{n,j,k} = \beta_{n,j,k}$ for $n = 1, 2, ..., 5$. Thus, the model we wish to estimate is

$$\begin{aligned} \ln E_{i,j,k,t} &= \alpha_{i,j,k} + \phi_{j,k,t} + \beta_{1,j,k} \ln p_{i,j,k,t} + \beta_{2,j,k} \ln pop_{i,t} + \beta_{3,j,k} \ln Y_{i,t} \\ &+ \beta_{4,j,k} y_{i,t} + \beta_{5,j,k} hdd^{2}_{i,t} + \delta_{j,k} K_{i,t} + \left(1-\gamma_{j,k}\right) \ln E_{i,j,k,t-1} + u_{i,j,k,t} \end{aligned} \qquad \text{(II-16)}$$

where $\alpha_{i,j,k}$, the country-specific effect for each commodity-sector, may be treated as random or fixed.

e. Data

Energy consumption statistics for the majority of our sample were collected from the International Energy Agency's (IEA) energy balances for both OECD and non-OECD countries. The dataset was checked for accuracy and consistency against a variety of

[43] The pooled sample mean of the deviation from the long run GDP growth rate, $K_{i,t}$, is equal to 0.0003 with a standard deviation of 0.0340 for the 50 countries in our study from 1990 to 2004. Thus we cannot reject the hypothesis that the average deviation from the long-run GDP growth rate is equal to zero in our study.

sources including the *Asia-Pacific Economic Cooperation Energy Database* and *The Energy Statistics Database* published by the Latin American Energy Organization (OLADE).

Energy demand is defined as end-use consumption of commercial energy [44] by commodity and sector. As stated previously, the sectors are industrial, transportation and commercial, residential, and agriculture (RCA). Industrial demand also includes "non-energy use" to capture the demand for some petroleum products and petrochemical feedstock which are not consumed for the purpose of providing energy. The energy commodities are electricity and all other direct-use fuels. Direct-use fuels include coal, petroleum products, natural gas, geothermal energy, solar, and wind. For the transportation sector, however, we choose to model total energy consumption rather than electricity and direct-use fuel separately because the share of electricity in total energy consumption for transportation is less than 1.5% for each country in our sample.

As a measure of aggregate output, we use real GDP denominated in 2000 dollars adjusted for purchasing power parity. GDP, consumer prices indices, wholesale price indices and population statistics were collected from the *World Development Indicators* database published by the World Bank. Missing series and observations were collected from the United Nations' *Common Database*.

Temperature data was collected from the *Global Summary of the Day* published by the National Climate Data Center of the National Oceanic and Atmospheric Administration of the United States Department of Commerce. The daily average of the

[44] Commercial energy excludes combustible renewables and waste such as industrial waste, municipal waste and biomass. While the IEA does publish a data series for combustible renewables and waste, it warns that the data is incomplete and not comparable across countries. Therefore, we focus on the impact of economic growth on commercial energy, which in any case is a better reflection of internationally traded energy commodities.

high and low temperature in the capital city is used to compute annual total heating degree days for each country in our sample.

We collected energy prices for each energy commodity-sector for all 50 countries in our study. While energy price statistics for OECD members and Latin America are widely published, the search for the remainder of the energy price data was extensive. We used the *Energy Prices and Taxes* data series published by the IEA for OECD members. The energy price index for transportation is the real household price index for unleaded gasoline. The share of industrial consumption of electricity, the real industrial price of electricity and the real industrial price for total energy are used to compute the share-weighted price of direct-use fuels for industrial consumers. Similarly, the real share-weighted household price indices of electricity and total energy are used to calculate the real share-weighted prices of electricity and direct-use fuels in the RCA sector.

For each member of OLADE, the Latin American Energy Organization, *The Energy Statistics Database* contains detailed statistics of nominal energy prices by fuel and sector. The real energy price index for transportation is computed using the consumption-weighted average nominal prices of gasoline, kerosene, diesel oil and fuel oil deflated by the consumer price index. The real industrial price index for electricity is calculated from the nominal price of industrial electricity and the wholesale price index. The real industrial price of direct-use fuels is obtained from the consumption-weighted nominal prices of the remaining fuels (natural gas, coal, liquid petroleum gases, diesel oil, and fuel oil) deflated by the wholesale price index. A similar methodology is applied to obtain the nominal price indices for electricity and direct-use fuels for each of the three components of the RCA sector. Then the consumption-weighted average nominal prices of electricity and direct-use fuels are deflated by the consumer price index to obtain the real RCA price indices of electricity and direct-use fuels.

The collection of price statistics was far more complicated for the remainder of our sample. These countries include China, Egypt, India, Indonesia, Israel, Malaysia, Pakistan, the Philippines, South Africa, and Thailand. Using a variety of sources, we extracted the real price of electricity and direct-use fuels from the components of consumer and wholesale price reports. Sources included national statistical websites, annual statistical yearbooks, the United Nations' *Statistical Yearbook for Asia and the Pacific*, and the International Labor Organization's *Laborsta Internet* database.

f. Model Specification: Endogenous Regressors

Recall equation (II-16), which describes the dynamic consumption model we wish to estimate for the five energy commodity-sector combinations we have defined.

$$\begin{aligned} \ln E_{i,j,k,t} &= \alpha_{i,j,k} + \phi_{j,k,t} + \beta_{1,j,k} \ln p_{i,j,k,t} + \beta_{2,j,k} \ln pop_{i,t} + \beta_{3,j,k} \ln Y_{i,t} \\ &+ \beta_{4,j,k} y_{i,t} + \beta_{5,j,k} hdd_{i,t}^2 + \delta_{j,k} K_{i,t} + \left(1-\gamma_{j,k}\right) \ln E_{i,j,k,t-1} + u_{i,j,k,t} \end{aligned} \tag{II-17}$$

In a panel setting, model estimation is complicated by the presence of the lagged dependent variable. Obviously, energy consumption ($\ln E_{i,j,k,t}$) is function of the country effect ($\alpha_{i,j,k}$) in this specification. Thus, the lagged consumption variable ($\ln E_{i,j,k,t-1}$) is also a function of the country effect ($\alpha_{i,j,k}$). This causes the standard OLS-based fixed or random effects estimator to be biased and inconsistent.[45]

[45] For a thorough discussion of panel data estimation see Baltagi (2008). To name just a few, see Greene (2003), Anderson and Hsiao (1981, 1982), Arellano and Bond (1991), Arellano and Bover (1995) and Ahn and Schmidt (1995) for a discussion of dynamic panel data.

We choose to estimate equation (II-16) using two-stage least squares (2SLS) as suggested by Balestra and Nerlove (1966).[46] The instruments are current and lagged values of the regressors. We do not assume, however, that lagged energy consumption is endogenous or that the remainder of the regressor set is exogenous. Instead, we estimate the model using the 2SLS-within estimator assuming one-way (country-specific) fixed effects [47] and perform the Durbin-Wu-Hausman (DWH) test for endogeneity of each regressor. The null hypothesis is that the OLS-within estimator of the model is consistent. The DWH test is essentially a Hausman test in which the covariance matrices for both 2SLS-within and the OLS-within estimator are based on the estimated error variance of the OLS-within estimator.

Lagged Energy Demand

First, we estimate the model assuming only lagged consumption, $\ln E_{i,j,k,t-1}$, is endogenous. The critical value of the DWH test, distributed as $\chi^2(1)$, is 3.84 at the 5% level and 6.63 at the 1% level. The results are presented in Table II-2 under Specification A. The test statistic is 2.12 for the transportation sector, 7.33 for industrial electricity, 17.75 for industrial direct-use, 8.25 for RCA electricity and 0.42 RCA direct-use fuels. Thus we reject the null hypothesis that we can consistently estimate our model using the OLS-within estimator in favor of 2SLS-within for three of the five commodity-sector combinations. Further, for consistency across the five models and given both empirical

---

[46] We chose to use the approach by Balestra and Nerlove (1966) to be consistent with Medlock and Soligo (2001). Alternatively, we could have used the procedure suggested by Arellano and Bond (1991). In this specification, the country-specific heterogeneity is removed by first differencing the model and estimating the parameters using a GMM estimator. The instrument set is composed of lagged levels of the endogenous variables and first differences of the exogenous variables. The system GMM is another alternative suggested by Blundell and Bond (1998).

[47] Further testing, described below, suggests that the assumption of a one-way fixed country effects model is appropriate.

and theoretical evidence, we will maintain the assumption throughout the rest of the paper that lagged energy demand is endogenous in each commodity-sector specification.

Table II-2: Durbin-Wu-Hausman test of the exogeneity of regressors

Specification:	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Regressor:	$\ln E_{i,j,k,t}$	$\ln Y_{i,t}$	$\ln p_{i,j,k,t}$	$y_{i,t}$	$\ln Pop_{i,t}$	$hdd^2_{i,t}$	$K_{i,t}$
Transportation	2.12	4.77**	1.21	0.34	9.13***	0.45	6.46**
Ind Electricity	7.33***	4.44**	0.12	0.30	3.87**	0.02	3.48
Ind Direct-use	17.75***	1.67	1.14	1.67	2.60	0.89	0.35
RCA Electricity	8.25***	0.69	3.30	0.02	1.02	10.49***	1.33
RCA Direct-use	0.42	0.23	2.06	2.23	7.51***	32.99***	0.30

The Durbin-Wu-Hausman test is distributed as $\chi^2(1)$. The null hypothesis is that the OLS-within estimator is consistent. Rejection of the null implies the regressor is endogenous to energy consumption.
** Denotes the statistic is significant at the 5% level. The critical value is 3.84.
*** Denotes the statistic is significant at the 1% level. The critical value is 6.63.

Total GDP

The exogeneity of the remaining regressors is also tested. The results are presented in Table II-3 under Specifications B through G. In each specification, only one additional variable is added to the set of endogenous regressors (and removed from the set of exogenous regressors). For example, in Specification B, total GDP ($\ln Y_{i,t}$) is included with lagged energy consumption to form the set of endogenous regressors. The DWH test statistics are reported in Table II-2. Also distributed as $\chi^2(1)$, the null hypothesis that total income is exogenous in our model of energy consumption is rejected at the 1% level for none of the commodity-sectors and 5% in only the transportation sector and industrial electricity sector. The test statistics are 4.77 for transportation, 4.44 for industrial electricity, 1.67 for industrial direct-use fuel, 0.69 for RCA electricity, and 0.23 for RCA direct-use fuel. Given only weak evidence of endogeneity, we opt to assume log GDP is exogenous in our model of energy demand to maintain symmetry across the five commodity-sectors.[48]

Price

In Specification C, the set of endogenous regressors is price ($\ln p_{i,j,k,t}$) and lagged energy demand while the set of exogenous regressors is total GDP, GDP per capita, population, and our parameters for the capital utilization and heating degree days squared. In no energy commodity-sector can we reject the hypothesis that price is exogenous with respect to contemporaneous consumption.

While it is fairly standard to treat price as exogenous in a model of energy demand, it need not be the case, particularly in our specification of energy consumption. Specifically, there is a transparent international market for the majority of direct-use fuels in which energy consumers are likely to be price takers. However, due to transmission constraints, electricity markets tend to be highly regionalized even within one nation's borders. For example, cold weather could result in an increase in electricity use as space heating demand increases. In a deregulated market, prices will increase to signal generators to produce more power. Thus, consumers may not be strictly price-takers, and the assumption of price exogeneity may be erroneous.

The inability to reject the hypothesis of price exogeneity in our model is likely a result of the unique aspects of global energy markets, as well as the use of yearly data and a disaggregated specification of energy demand. First, there are many inefficiencies in

---

[48] Empirical evidence on the exogeneity of income is mixed. Darrat, Gilley, and Meyer (1996), Stern (1993, 2000), and Moroney (1992) study the causal relationship between total energy consumption and income. More recently, the causal relationship between income and electricity consumption has been analyzed by authors such as Wolde-Rufael (2006) and Ferguson, Wilkenson, and Hill (2006) which has also yielded mixed results.

energy markets world-wide and across commodities which may hinder the ability of energy prices to respond to demand. For example, national policies designed to promote economic development may support fuel and electricity subsidies, fixed prices and price caps. Also, the global crude oil market is dominated by a cartel, the Organization of Petroleum Exporting Countries (OPEC), which controls roughly 40% of the world's crude oil production. According to OPEC, it "seeks to ensure the stabilization of oil prices in international oil markets, with a view to eliminating harmful and unnecessary fluctuations." Production quotas are the organization's primary tool in its battle against price volatility. The following quote from OPEC alone provides support to the notion that price is exogenous to energy demand. "If demand grows, or some producers supply less oil, OPEC can increase its oil production to prevent a sudden rise in prices or shortfall in supply."[49]

Second, the use of yearly data likely masks the impact of short term events on demand, as well as price. For example, consider the U.S. natural gas market which has a well-developed, transparent daily spot market. On extremely cold winter days, spot prices rise sharply in response to increased demand for home heating. However, when price and demand are averaged over the entire year, these short term events may no longer be significant.

Finally, the disaggregation of energy demand into five commodity-sector combinations, also likely lends support to the exogeneity of price on energy demand. As discussed previously, electricity consumption is a derived demand which likely differs substantially between industrial users and residential (or RCA) users. Thus, one sector may have little impact on the overall electricity market. It is interesting to note, however,

[49] *Frequently Asked Questions*, Organization of the Petroleum Exporting Countries, March 2009.

that when we estimate the same model using Specification C for total energy demand and total commodity-sector weighted prices, the DHW statistic is 7.88. One can safely reject the null hypothesis at the 1% level, suggesting that energy prices are indeed a function of aggregate energy demand. Nonetheless, we proceed assuming price exogeneity in the five models of energy commodity-sector demand.

GDP per Capita

In Specification D, the set of endogenous regressors is composed of GDP per capita ($y_{i,t}$) and lagged energy demand while the exogenous regressors are total GDP price, population, and our parameters for capital inefficiency and heating degree days squared. Recall the DWH test statistic is distributed as $\chi^2(1)$, the 5% critical value is 3.83 and the 1% critical value is 6.63. Our tests yields statistics equal to 0.34 for transportation, 0.30 for industrial electricity, 1.67 for industrial direct-use fuel, 0.02 for RCA electricity and 2.23 for RCA direct-use fuel. Thus, we can not reject the hypothesis that income per capita is exogenous in the five commodity-sectors.

Population

The exogeneity of population is tested in Specification E. While there seems to be no economic justification for population to be a function of energy demand, the DWH test statistic for endogeneity is significant at the 1% level for transportation and 5% level for RCA direct-use. Specifically, the test statistics are 9.13 for transportation, 3.87 for industrial electricity, 2.60 for industrial direct-use fuel, 1.02 for RCA electricity and 7.51 for RCA direct-use fuel. Because we find that the test statistic is insignificant for total energy demand and want to maintain consistency across all commodity-sectors, we chose to assume population is exogenous.

Squared Heating Degree Days

The exogeneity of the last regressor, squared heating degree days ($hdd_{i,t}^2$), is tested in Specification F. While the endogeneity of squared heating degree days can not be rejected in either RCA electricity or RCA direct-use, it can be rejected in the remaining three commodity-sectors. The DWH test statistic is 0.45 for transportation, 0.02 for industrial electricity, 0.89 for industrial direct-use fuels, 10.49 for RCA electricity and 32.99 for RCA direct-use fuels. Recall that heating degree days are used to capture demand for space heating, which is the result of cold weather. Because there is no plausible theory to support this endogeneity in the RCA sector, we chose to assume squared heating degree days are exogenous.

Capital Utilization

Finally recall the definition of the change in capital utilization variable which is defined as the deviation from the long run GDP growth rate in equation (II-13).

$$K_{i,t} = \left(\ln Y_{i,t} - \ln Y_{i,t-1}\right) - \frac{1}{10}\sum_{s=t-10}^{t-1}\left(\ln Y_{i,s} - \ln Y_{i,s-1}\right)$$

In Specification G, we test the exogeneity of $K_{i,t}$. The DWH test statistic is 6.46 for transportation, 3.48 for industrial electricity, 0.35 for industrial direct-use fuels, 1.33 for RCA electricity and 0.30 for RCA direct-use fuels. Thus at the 1% level, we can not reject the null hypothesis of exogeneity in any of the five commodity-sector combinations. Although the DWH test statistic is significant at the 5% level in the transportation sector, to maintain a similar model across all five energy commodity-sector the change in capital utilization is assumed to be exogenous throughout the rest of this paper.

g. Model Specification: Country and Time Effects

It remains to test the significance and nature of the effects of the explanatory variables in our model of energy demand. We maintain the assumption that lagged demand is the only endogenous regressor. In the analysis to follow, we use 2SLS-panel estimators where the instruments are present and lagged values of the regressors and rely heavily on the F-test. Wooldridge (1990) points out that the distribution of the F-statistic for the 2SLS estimator is unknown, even asymptotically. Thus, inferences made from the standard F-statistic can be misleading. Instead, Wooldridge suggests using the sum of squared residuals from the second-stage regression for both the restricted and unrestricted estimations in the numerator. The denominator remains the residuals from the 2SLS estimation. We follow this methodology in each of the F-tests below.

Joint Significance of Country and Time Effects

For our model, we first test the joint significance of the country and time effects in equation (II-16) above. The unrestricted estimation is a two-way 2SLS-within estimator where the intercept is allowed to vary by year and country while the slopes for the exogenous variables are assumed to be constant. The restricted model uses pooled 2SLS which assumes the slopes and intercept are homogeneous across all countries. This amounts to a 2SLS F-test of the null hypothesis that both the country and time effects are equal to zero, i.e. $\alpha_{i,j,k} = \phi_{j,k,t} = 0$. Letting N, T and K respectively equal the number of countries, years and regressors (excluding the constant) in our model, the number of restrictions equals $(N-1)+(T-1)$, and the number of degrees of freedom is $(N-1)(T-1)-K$. Recall that we have a panel of 50 countries across 15 years and that there are 7 regressors. Therefore, the 2SLS F-test for joint fixed effects is distributed as $F(63,679)$ where the 5% critical value equals 1.33 and the 1% critical value equals 1.49. The results, presented in Table II-3, show that the null hypothesis of zero country and time effects is strongly rejected in each of the five energy commodity-sector specifications. Specifically, the 2SLS F-statistic equals 3.96 for transportation, 4.50 for industrial electricity, 2.17 for industrial direct-use fuels, 5.05 for RCA electricity and 2.54 for RCA direct-use fuels.

Table II-3: 2SLS F-test for the significance of effects

Null Hypothesis:	$\alpha_{i,j,k}=0,\phi_{j,k,t}=0$	$\phi_{j,k,t}=0$	$\alpha_{i,j,k}=0$
Distribution:	$F(63,679)$	$F(14,679)$	$F(49,693)$
CV 5%:	1.33	1.71	1.37
CV 1%:	1.49	2.11	1.56
Transportation	3.96***	1.37	4.65***
Industrial Electricity	4.50***	1.39	5.34***
Industry Direct-use	2.17***	1.55	2.30***
RCA Electricity	5.05***	0.54	6.42***
RCA Direct-use	2.54***	1.42	2.83***

Note that the computation of the F-statistic for 2SLS requires the use of the sum of squared residuals from the second-stage regression for both the restricted and unrestricted estimation for the numerator. The denominator, however, uses the sum of squared residuals from the 2SLS regression for the unrestricted model. See Wooldridge (1990).
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

Significance of Time Effects

Next we test the assumption that time effects are equal to zero while still allowing for country-specific fixed effects. The unrestricted estimation remains the two-way 2SLS-within estimator where the slopes are assumed to be constant and the intercept is allowed to vary by year. However, the restricted model becomes the one-way 2SLS-within estimator where the slopes are common but the intercept may vary by country but not by year. Although the number of degrees of freedom remains unchanged, there are now $(T-1)$ restrictions. The F-statistic for this test is distributed as $F(14,679)$ with a 5% critical value of 1.71 and a 1% critical value of 2.11. As shown in Table II-3, we can not reject the null hypothesis that the time effects are equal to zero, $\phi_{j,k,t}=0$, at even the 5% level in any of the five energy commodity-sectors. The 2SLS F-statistic is 1.37 for transportation, 1.39 for industrial electricity, 1.55 for industrial direct-use fuels, 0.54 for RCA electricity and 1.42 for RCA direct-use fuels. We therefore conclude that it is reasonable to exclude time effects from our model of energy demand.

Significance of Country Effects

Because we strongly rejected the hypothesis that country and time effects are jointly equal to zero but not the hypothesis that time effects are equal to zero, one may reasonably conclude that the country effects must be significant. We demonstrate this with a third set of 2SLS F-tests, which are presented in the last column of Table II-3. Here the unrestricted model is the one-way 2SLS-within estimator allowing for only country-specific heterogeneity in the intercept. Again, the slopes on the exogenous variables are assumed to be common. The restricted model is the pooled-2SLS model with a common intercept and common slopes. There are now $(N-1)$ restrictions and $N(T-1)-K$ degrees of freedom. Note that the number of degrees of freedom has increased from 679 to 693 because we no longer need to estimate the time effect parameters. The critical values for an F-statistic distributed as $F(49,693)$ are 1.37 at the

5% level and 1.56 at the 1% level. We easily reject the null hypothesis in favor of the one-way 2SLS-within estimator for all five energy commodity-sectors. The 2SLS F-statistic is equal to 4.65 for transportation, 5.34 for industrial electricity, 2.30 for industrial direct-use fuels, 6.34 for RCA electricity and 2.83 for RCA direct-use fuels.

Fixed vs. Random Country Effects

Finally, the one-way country effects may be treated as either fixed or random. The fixed effects model allows the effect to be correlated with the regressors and is always consistent. However, if the country effects are strictly uncorrelated with the regressors, the random effect estimator is consistent and efficient since it decreases the number of parameters to be estimated by (N-1). Thus, we perform a Hausman-type test to compare the estimated coefficients. Under the null hypothesis, the effects are uncorrelated with the regressors. Therefore, there will be no systematic difference between the fixed and random effects parameter estimates since each will converge to its true value.

Table II-4: Hausman test for random effects

Null Hypothesis:	2SLS-GLS is consistent
Distribution:	$\chi^2(7)$
CV 5%:	14.1
CV 1%:	18.5
Transportation	61.2***
Industrial Electricity	79.2***
Industry Direct-use	43.0***
RCA Electricity	421.4***
RCA Direct-use	96.8***

** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

For this Hausman test, the consistent estimator is 2SLS-within while the efficient (and consistent under the null hypothesis) estimator is 2SLS-GLS. The test is distributed as $\chi^2(7)$ with critical values equal to 14.1 at the 5% level and 18.5 at the 1% level. The results are presented in Table II-4. For each of the five energy commodity-sectors, we easily reject the 2SLS-GLS estimator in favor of the 2SLS-within estimator. The test statistic is 61.2 for transportation, 79.2 for industrial electricity, 43.0 for industrial direct-use, 421.4 for RCA electricity and 96.8 for RCA direct-use.

h. Model Specification: Error Covariance Matrix

Consider a basic fixed effects model, $y_{i,t} = x_{i,t}\beta + v_{i,t}$ where the composite error is $v_{i,t} = a_i + u_{i,t}$. Efficient estimation of the 2SLS-within model requires the idiosyncratic errors, $u_{i,t}$, to be homoskedastic and serially correlated. Speficially, $E\left(u_i u_i'\right) = \sigma_u^2 \mathrm{I}_T$, and it can be shown (see Wooldridge (2002) or Baltagi (2008), for example) that under this condition, the serial correlation takes a special form, namely, $corr\left(u_{i,t}, u_{i,s}\right) = -1/(T-1)$ for all $t \neq s$.

As suggested by Wooldridge (2002), we test our model to determine if the estimated idiosyncratic errors exhibit this form of serial correlation. Using the predicted errors of our model (purged of the fixed effect), we perform a pooled OLS regression of $\hat{u}_{i,j,k,t}$ on $\hat{u}_{i,j,k,t-1}$ using the Huber-White robust estimates of the standard errors. The null hypothesis is that the coefficient on $\hat{u}_{i,j,k,t-1}$ equals $corr\left(u_{i,j,k,t}, u_{i,j,k,t-1}\right)$ or -0.0714, in the case of our model. The test is distributed as $F(1,698)$ with critical values equal to 3.85 at the 5% level and 6.67 at the 1% level. For all sectors and commodities, we are unable reject the null that $corr\left(u_{i,j,k,t},u_{i,j,k,t-1}\right)=-1/(T-1)$ at the 1% level. At the 5% level, this test is rejected only for industrial electricity demand. Table II-5 contains the resulting F-statistics for this test. Specifically, the test statistics are 0.01 for transportation, 5.80 for industrial electricity, 0.79 for industrial direct-use fuels, 1.18 for RCA electricity, and 0.03 for RCA direct-use fuels.

Table II-5: Tests for serial correlation and homoskedasticity

Null Hypothesis:	$corr\left(u_{i,j,k,t},u_{i,j,k,t-1}\right)=-1/(T-1)$	$\sigma_{u,i}^2=\sigma_u^2$ for all $i=1,2,\ldots,N$
Distribution:	$F(1,698)$	$\chi^2(50)$
CV 5%:	3.85	67.5
CV 1%:	6.67	76.2
Transportation	0.01	6,101.4***
Industrial Electricity	5.80**	22,356.3***
Industry Direct-use	0.79	15,910.0***
RCA Electricity	1.18	3,416.7***
RCA Direct-use	0.03	3,697.5***

** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

This test, however, does not preclude heteroskedasticity. We, therefore, perform a modified Wald test for the presence of country-specific heterogeneity of the error variances, as suggested by Greene (2003). The null hypothesis is $\sigma_{u,i}^2=\sigma_u^2$ for all $i=1,2,\ldots,50$. The critical values for this test, which is distributed as $\chi^2(50)$, are 67.5 at the 5% level and 76.2 at the 1% level. The results are presented in the last column of Table II-5. Notice we reject the null hypothesis of homoskedasticity in favor of country-specific heterogeneity of the error variances for all energy commodity-sector combinations. The test statistics are 6101 for transportation, 22356 for industrial electricity, 15910 for industrial direct-use fuels, 3417 for RCA electricity and 3678 for RCA direct-use fuels.

Given the strong evidence of heteroskedasticity of the idiosyncratic errors, we opt to use the 2SLS-within robust variance estimator suggested by Arellano (1987). It should be noted that altering the specification of the standard errors has no impact on the parameter estimates of the coefficients. However, the use of robust standard errors will yield more conservative estimates of standard errors and parameter significance.

i. Model Estimation and Simulated Demand Paths

The 2SLS-within parameter estimates and robust standard errors for the short-run coefficients in our model of energy consumption (equation (II-16)) are presented in Table II-6. Recall previous tests indicate that we cannot reject the hypothesis that the time effects, $\phi_{j,k,t}$, are all jointly equal to zero in each commodity-sector. Therefore, the time effects are omitted from the estimation of our model. The implied coefficients of the long-run model (equation (II-7)) are presented below the short-run parameter estimates. Finally, if $b_3 > 0$ and $b_4 < 0$, then the income level at which the long run income elasticity equals zero, i.e. $\varepsilon^{Y*} = 0$, maximizes demand. This is denoted $y^{*\max}$ in Table II-6.

In general, all parameter estimates significantly different than zero have the appropriate sign. Moreover, patterns of significance across sectors and commodities lend support to disaggregating total energy consumption into the demand for transportation fuels, industrial electricity, industrial direct-use fuels, RCA electricity and RCA direct-use fuels. For example, price elasticity appears highest in the transportation sector and for electricity in general while elasticity with respect to temperature (implied by the parameter on heating degree days squared) appears highest in the RCA sector.

Table II-6: Estimation results for the five energy commodity-sectors

Short-run energy demand model (equation (II-16) excluding time effects):

$$\ln E_{i,j,k,t} = \alpha_{i,j,k} + \beta_{1,j,k} \ln p_{i,j,k,t} + \beta_{2,j,k} \ln pop_{i,t} + \beta_{3,j,k} \ln Y_{i,t} + \beta_{4,j,k} y_{i,t} + \beta_{5,j,k} hdd^2_{i,t} + \delta_{j,k} K_{i,t} + \left(1-\gamma_{j,k}\right) \ln E_{i,j,k,t-1} + u_{i,j,k,t}$$

Long-run energy demand model (equation (II-7) excluding time effects):

$$\ln E^*_{i,j,k,t} = a_{i,j,k} + b_{1,j,k} \ln p_{i,j,k,t} + b_{2,j,k} \ln pop_{i,t} + b_{3,j,k} \ln Y_{i,t} + b_{4,j,k} y_{i,t} + b_{5,j,k} hdd^2_{i,t}$$

Sector: Commodity:	Transport Total	Industrial Electricity	Industrial Direct-Use	RCA Electricity	RCA Direct-Use
Estimated short-run coefficients					
$\ln p_{i,j,k,t}$	-0.0973*** (0.0148)	-0.0536*** (0.0181)	-0.0250 (0.0237)	-0.0624*** (0.0087)	-0.0069 (0.0193)
$\ln pop_{i,t}$	0.0398 (0.0704)	-0.0404 (0.1302)	0.0472 (0.1927)	0.0975 (0.0759)	0.2933** (0.1286)
$\ln Y_{i,t}$	0.2426*** (0.0602)	0.2672*** (0.0543)	0.1887* (0.1073)	0.2037*** (0.0514)	0.0372 (0.0456)
$y_{i,t}$	-0.0018 (0.0020)	-0.0078*** (0.0026)	-0.0047 (0.0049)	-0.0040* (0.0024)	0.0008 (0.0025)
$hdd^2_{i,t}$	0.0008 (0.0005)	-0.0006 (0.0005)	0.0014** (0.0006)	0.0017*** (0.0006)	0.0053*** (0.0008)
$K_{i,t}$	0.4261*** (0.0571)	0.5276*** (0.1000)	0.5561*** (0.1209)	0.0219 (0.0558)	0.1873** (0.0889)
$\ln E_{i,j,k,t-1}$	0.7698*** (0.0459)	0.8103*** (0.0473)	0.7719*** (0.0520)	0.8281*** (0.0221)	0.7270*** (0.0400)
Implied long-run coefficients					
$\ln p_{i,j,k,t}$	-0.4226	-0.2840	-0.1098	-0.3631	-0.0252
$\ln pop_{i,t}$	0.1729	-0.2129	0.2070	0.5670	1.0745
$\ln Y_{i,t}$	1.0540	1.4082	0.8245	1.2059	0.1365
$y_{i,t}$	-0.0078	-0.0414	-0.0204	-0.0236	0.0031
$hdd^2_{i,t}$	0.0034	-0.0031	0.0060	0.0101	0.0193
$y^{*\max}$	135,522***	34,038***	40,444***	51,067***	n/a
Within R^2	0.9395	0.9207	0.6961	0.9711	0.7414

Recall regressors $y_{i,t}$ and $hdd^2_{i,t}$ are income per capita defided by 1000 and heating degree days divided by 1000 then squared, repectively. Reported significance of $y^{*\max}$ is with respect to the joint significance of total GDP, GDP per capita and capital utilization.

*, **, *** Denote the statistic is significant at the 10%, %5 or 1% level, respectively.

We test the appropriateness of disaggregating total energy consumption into the five energy commodity-sectors using a 2SLS-within F-test of equality of the parameters in the five models. The results are presented at the top of Table II-7. Under the null hypothesis, the seven parameter estimates are individually and simultaneously equal, meaning our five energy commodity sectors can be pooled into one model of total energy consumption. Note that we are *not* testing that the seven coefficients are equal to a single value, rather each of the seven coefficients is equal to a single value across the five models. A rejection of the null hypothesis implies there are systematic differences in the demand for energy across the energy commodity-sectors. Distributed as $F(28,249)$, the 2SLS-within F-statistic is equal to 7.23. With a critical value of 1.80, we reject the hypothesis at the 1% level. We conclude the five models of energy consumption by commodity and sector cannot be pooled into a single model of total energy consumption.

Table II-7: Poolability of commodity-sectors into one model of total energy demand

Parameters:	Null Hypothesis:	Distribution:	CV 5%:	CV 1%:	F-stat:
All	$\beta_{n,j,k}=\beta_n\ ,\forall j,k,n$ $\delta_{j,t}=\delta\ \ \forall j,k$, $(1-\gamma_{j,k})=(1-\gamma)\ ,\forall j,k$	$F(28,249)$	1.52	1.80	7.23***
$\ln Y, y, K$	$\beta_{n,j,k}=\beta_n\ ,\forall j,k,n$	$F(12,249)$	1.79	2.25	8.49***
$\ln p$	$\beta_{1,j,k}=\beta_1\ \ \forall j,k$	$F(4,249)$	2.41	3.40	4.14***
$\ln pop$	$\beta_{2,j,k}=\beta_2\ \forall j,k$	$F(4,249)$	2.41	3.40	1.02
$\ln Y$	$\beta_{3,j,k}=\beta_3\ \ \forall j,k$	$F(4,249)$	2.41	3.40	3.47***
y	$\beta_{4,j,k}=\beta_4\ \ \forall j,k$	$F(4,249)$	2.41	3.40	1.64
hdd^2	$\beta_{5,j,k}=\beta_5\ \forall j,k$	$F(4,249)$	2.41	3.40	10.47***
K	$\delta_{j,k}=\delta\ \forall j,k$	$F(4,249)$	2.41	3.40	10.24***
$\ln E_{t-1}$	$(1-\gamma_{j,k})=(1-\gamma)\ ,\forall j,k$	$F(4,249)$	2.41	3.40	1.44

* Denotes the statistic is significant at the 10% level.
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

Next, in an effort to understand how demand in the five energy commodity-sectors differs, we test the equality of each parameter estimate individually but not simultaneously. The details are presented in Table II-7. The seven 2SLS-within F-tests are distributed as $F(4,249)$. The critical value is equal to 2.41 at the 1% level and 3.40 at the 5% level.

In the first of these tests, the F-statistic for the equality of the coefficient on lagged consumption is equal to 1.44. Therefore, the lagged nature of energy consumption does not appear to vary statistically by commodity-sector. Moreover, while $\left(1-\gamma_{j,k}\right)$ is statistically less than 1 for each energy commodity-sector, the result that $\left(1-\gamma_{j,k}\right)>0.7$ suggests the majority of contemporaneous demand is a function of lagged demand (i.e. capital stock) rather than the long run equilibrium implied by contemporaneous factors.

The coefficients on population are also statistically equal across the five models of energy commodity-sector consumption.[50] The 2SLS-within F-statistic for joint equality is equal 1.03. As shown in Table II-6, RCA direct-use fuel demand is the only commodity-sector in which population is statistically significant and greater than zero.

The F-statistic for the equality of the coefficient on price in all five commodity-sectors is equal to 4.14, and we conclude that price elasticity varies statistically in the five

[50] We also estimated a per capita energy demand model, similar to the model specified by Medlock and Soligo (2001) with unsatisfactory results. The 2SLS-within per capita model we estimated was

$$\ln ec_{i,t} = \alpha_{i,j,k,t} + \beta_{1,j,k} \ln p_{i,j,k,t} + \beta_2 \ln y_{i,t} + \beta_3 \left(\ln y_{i,t}\right)^2 + \beta_4 hdd_{i,t}^2 + \delta_{j,k} K_{i,t} + \left(1-\gamma_{j,k}\right) \ln ec_{i,j,k,t-1}$$

In this specification, if $\beta_{2,j,k} > 0$ and $\beta_{3,j,k} < 0$ then the energy demand is maximized at an income level equal to $y_{j,k}^{*} = \exp\left(\beta_{2,j,k} / -2\beta_{3,j,k}\right)$. Although $\beta_{2,j,k}$ was found to be at least weakly significant in each energy commodity-sector model, $\beta_{3,j,k}$ was significantly less than zero in only the per capita industrial electricity demand model. However, we did find $\delta_{j,k}$ to be significantly greater than zero which supports the inclusion of capital utilization parameter in our model.

commodity-sectors. The short-run coefficient on price, shown in Table II-6, is significantly less than zero in transportation energy use and industrial and RCA electricity. However, the parameter estimate is not significant in the demand for direct-use fuel in either the industrial or RCA sector. Recall that in these sectors a portion of demand is derived from space heating. These results suggest that the demand for heating is inelastic with respect to price. Demand in the industrial sector is also derived from using direct-use fuel in the production of goods. If the input price of energy is passed on to the consumers of the finished products, industrial producers may be indifferent to the price of direct-use fuels.

Furthermore, the demand derived from space-heating is not equal across the five commodity-sectors. As shown in Table II-7, the F-statistic is equal to 10.47. In Table II-6, the short run coefficient on heating degree days squared is not significant for either transportation fuel use or industrial electricity consumption. The coefficient on heating degree days squared, however, is significantly greater than zero at the 1% level for both electricity and direct-use fuel consumption in the RCA sector and at the 5% level for direct-use fuels in the industrial sector. Holding all else constant, we conclude that energy consumption – and likely annual variation –will be greater in these three sectors for cold climates nations.

It remains to test the equality of the three income variables – total GDP, GDP per capita and the capital utilization. As discussed above, we found the rate of adjustment to the long run equilibrium to be statistically equal in the five models. However, short run deviations from the long run equilibrium are also a function of capital utilization in our model. We reject the hypothesis that $\delta_{j,k}$ is equal in each energy commodity-sector at the 1% level with a 2SLS-with F-statistic equal to 10.24. The estimates of $\delta_{j,k}$ are highest in the industrial sector for both commodities, followed by the transportation sector and direct-use fuel demand in the RCA sector. However, $\delta_{j,k}$ is insignificant for electricity demand in the RCA sector. Therefore, deviations from the long run GDP growth rate have the biggest impact on energy intensity in the industrial and transportation sectors, respectively.

In the long and short run, total GDP and GDP per capita capture the relationship between energy demand and economic growth. Differences in this relationship will cause $\beta_{3,j,k}$ and $\beta_{4,j,k}$ to vary statistically by sector j and commodity k. The 2SLS-with F-statistics for the equality of the coefficients individually across all five models are 3.47 for total GDP and 1.64 for GDP per capital. Thus we reject the hypothesis that $\beta_{3,j,k}$ is equal in all five energy commodity-sectors but not $\beta_{4,j,k}$.

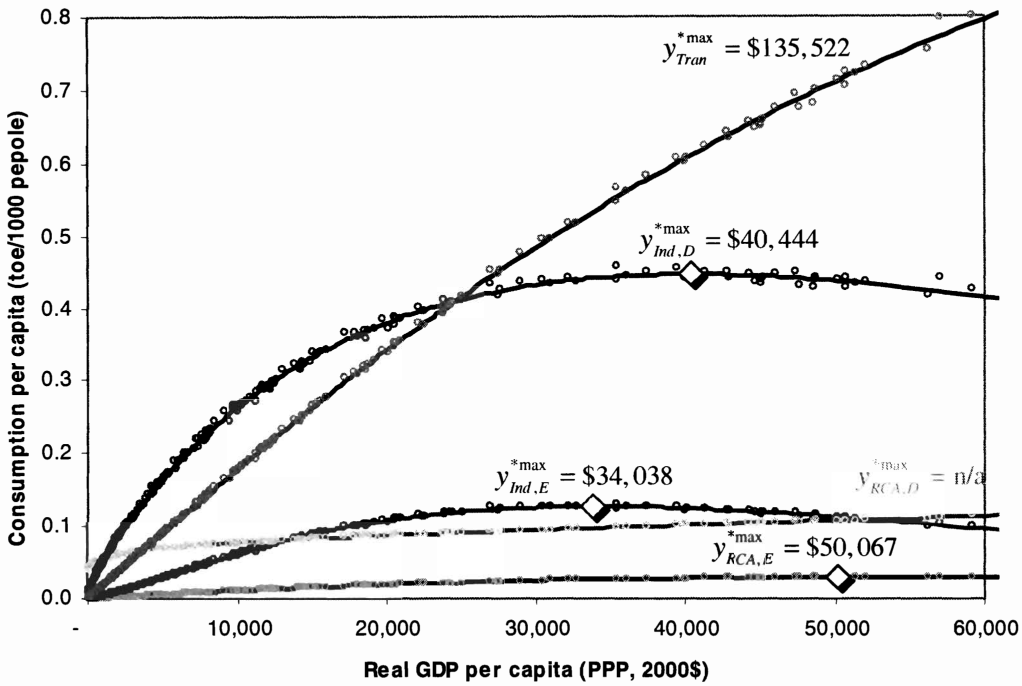
For the consumption of electricity, it can be shown that $b_{3,j,k} > 0$ and $b_{4,j,k} < 0$ in both the industrial and RCA sectors. Therefore, long run income elasticity of electricity demand is initially positive but decreases with GDP per capita, resulting in a non-monotonic relationship between economic development and electricity consumption. From equation (II-8) electricity demand is maximized at \$34,038 and \$51,067 in the long run in the industrial and RCA sectors, respectively.

Although the coefficient on GDP per capita in Table II-6 is not significant for transportation or direct-use fuel consumption in the industrial or RCA sectors, the point estimates are negative. From these point estimates, energy demand is maximized with respect to GDP per capita at \$135,522 in the transportation sector and \$40,444 in the industrial sector for direct-use fuel. The parameter estimate for GDP per capita is positive, although not statistically greater than zero, in RCA direct-use fuel consumption, and income elasticity is equal to zero at an income level equal to -$44,184. However, this would be the point at which demand is *minimized* with respect to income. Therefore, the evidence suggests income elasticity of RCA direct-use fuel consumption is not strictly decreasing over the relevant income range.

Finally, we test whether the three income variables are jointly equal across the five models of energy demand. The 2SLS-with F-test is distributed as $F(12, 249)$ and yields a test statistic equal to 8.49, which far exceeds the 1% critical value of 2.26. We therefore conclude that economic development impacts energy demand differently across the five energy sector-commodities.

To illustrate the varying relationship between economic development and energy demand in the five commodity-sectors, Figure II-13 plots simulated paths of per capita energy consumption by commodity-sector as a function of GDP per capita for a hypothetical country using the parameter estimates presented in Table II-6. The intercept for each commodity-sector is the sample average of the country-specific intercept. Annual heating degree days are assumed to equal the sample average of 4,274 in each year. Population and the real energy price index in each commodity-sector are also held constant at 1,000,000 people and 100 2000 dollars adjusted for purchasing power parity, respectively. For each-commodity sector, two simulated demand paths are presented. The first, shown as a solid line, represents the demand path where GDP per capita grows at a constant rate equal to the sample average of 2.15% per year. In contrast, the demand path shown in hollow circles includes the impact of capital utilization where GDP per capita is allowed to grow at a random rate with a mean of 2.15% and standard deviation of 3.38%.[51] The variation between the two scenarios represents the impact of changing capital utilization.

Figure II-13: Simulated energy consumption for a hypothetical country



As shown in Figure II-13, demand is maximized (and income elasticity equals zero) at per capita income levels of \$34,038 and \$40,444 respectively in the industrial sector for electricity and direct-fuel, which is lower than in the RCA and transportation sectors. Moreover, the impact of changing capital utilization appears greatest in transportation and industrial direct-use fuel demand. This is consistent with the results of presented in Table II-6 where $\delta_{Ind,k} > \delta_{Tran} > \delta_{RCA,k}$. The remainder of this section

[51] Note that total GDP grows at an anverage annual rate of 3.28% with a standard deviation of 3.49% in the pooled sample. The mean annual population growth rate is equal to 1.10% with a standard deviation of 0.84%.

compares the relationship between energy demand and economic development by sector and then by commodity.

Comparison of Demand by Sector

Equation (II-5) can be used to generate the three long run models of demand by sector. Industrial sector energy demand is simply the sum of industrial electricity and direct-use fuel consumption.

$$\begin{aligned} E^*_{Ind,i,t} &= E^*_{Ind,E,i,t} + E^*_{Ind,D,i,t} \\ &= A_{Ind,E,i,t}\, p_{Ind,E,i,t}^{b_{1,Ind,E}}\, pop_{i,t}^{b_{2,Ind,E}}\, Y_{i,t}^{b_{3,Ind,E}} \exp(y_{i,t})^{b_{4,Ind,E}} \exp(hdd_{i,t}^2)^{b_{5,Ind,E}} \\ &\quad + A_{Ind,D,i,t}\, p_{Ind,D,i,t}^{b_{1,Ind,D}}\, pop_{i,t}^{b_{2,Ind,D}}\, Y_{i,t}^{b_{3,Ind,D}} \exp(y_{i,t})^{b_{4,Ind,D}} \exp(hdd_{i,t}^2)^{b_{5,Ind,D}} \end{aligned} \tag{II-18}$$

Similarly, RCA energy demand is the sum of RCA consumption of electricity and direct-use fuel.

$$\begin{aligned} E^*_{RCA,i,t} &= E^*_{RCA,E,i,t} + E^*_{RCA,D,i,t} \\ &= A_{RCA,E,i,t}\, p_{RCA,E,i,t}^{b_{1,RCA,E}}\, pop_{i,t}^{b_{2,RCA,E}}\, Y_{i,t}^{b_{3,RCA,E}} \exp(y_{i,t})^{b_{4,RCA,E}} \exp(hdd_{i,t}^2)^{b_{5,RCA,E}} \\ &\quad + A_{RCA,D,i,t}\, p_{RCA,D,i,t}^{b_{1,RCA,D}}\, pop_{i,t}^{b_{2,RCA,D}}\, Y_{i,t}^{b_{3,RCA,D}} \exp(y_{i,t})^{b_{4,RCA,D}} \exp(hdd_{i,t}^2)^{b_{5,RCA,D}} \end{aligned} \tag{II-19}$$

Finally, transportation fuel demand follows directly from equation (II-5).

$$E^*_{Tran,i,t} = A_{Tran,i,t}\, p_{Tran,i,t}^{b_{1,Tran}}\, pop_{i,t}^{b_{2,Tran}}\, Y_{i,t}^{b_{3,Tran}} \exp(y_{i,t})^{b_{4,Tran}} \exp(hdd_{i,t}^2)^{b_{5,Tran}} \tag{II-20}$$

The long run income elasticity of transportation fuel demand is described by equation (II-8). However, the long run income elasticity of energy demand in the industrial sector is the share-weighted elasticity of demand by commodity. Specifically,

$$\varepsilon^{*Y}_{Ind,i,t} = \frac{E^*_{Ind,E,i,t}\varepsilon^{*Y}_{Ind,E,i,t} + E^*_{Ind,D,i,t}\varepsilon^{*Y}_{Ind,D,i,t}}{E^*_{Ind,i,t}}.$$

The income level at which the long run elasticity of demand equals zero is also a function of the commodity weights. Cancelling total industrial energy demand, $E^{*}_{Ind,i,t}$, from both the numerator and denominator,

$$y^{*\max}_{Ind,i,t} = \frac{-\left(E^{*}_{Ind,E,i,t}b_{3,Ind,E} + E^{*}_{Ind,D,i,t}b_{3,Ind,D}\right)}{E^{*}_{Ind,E,i,t}b_{4,Ind,E} + E^{*}_{Ind,D,i,t}b_{4,Ind,D}}.$$

Because $b_{3,Ind,E} > 0$, $b_{3,Ind,D} > 0$, $b_{4,Ind,E} < 0$ and $b_{4,Ind,D} < 0$ in Table II-6, $y^{*}_{Ind,i}$ maximizes the long run equilibrium demand for energy in the industrial sector. Moreover, because $y^{*\max}_{Ind,E} \leq y^{*\max}_{Ind,D}$, then $y^{*\max}_{Ind,i,t} \geq y^{*\max}_{Ind,E}$ and $y^{*\max}_{Ind,i,t} \leq y^{*\max}_{Ind,D}$. Thus, industrial demand is maximized between \$34,038 and \$40,444, and the higher the share of electricity demand in the industrial sector, the lower $y^{*\max}_{Ind,i,t}$ falls within this range.

Similarly, the long run income elasticity of RCA energy demand is equal to

$$\varepsilon^{*Y}_{RCA,i,t} = \frac{E^{*}_{RCA,E,i,t}\varepsilon^{*Y}_{RCA,E,i,t} + E^{*}_{RCA,D,i,t}\varepsilon^{*Y}_{RCA,D,i,t}}{E^{*}_{RCA,i,t}}$$

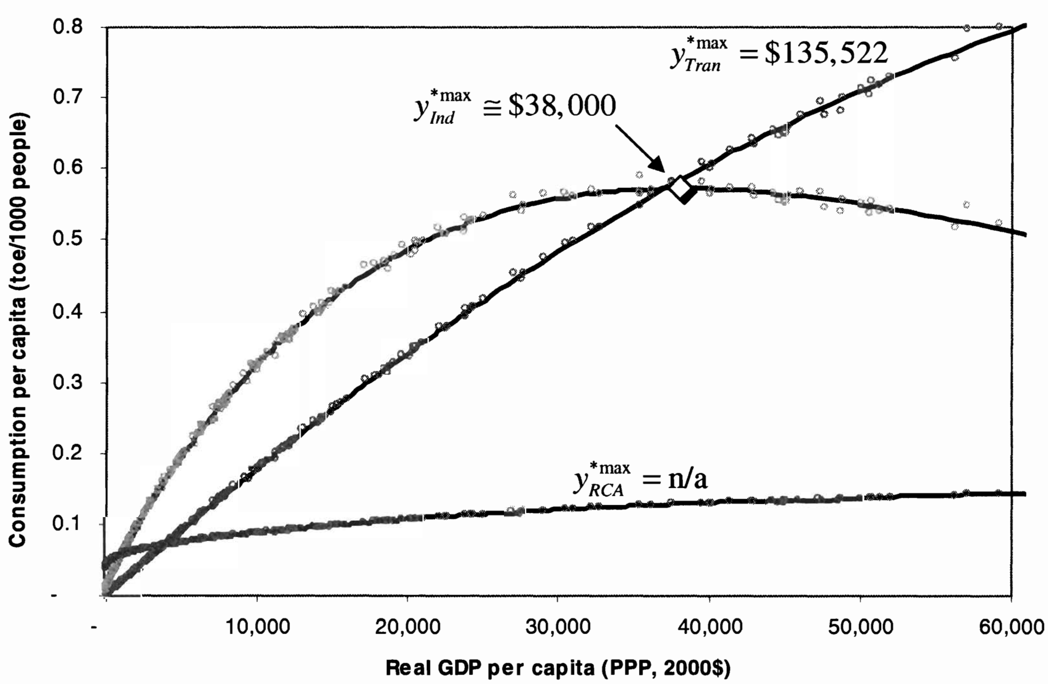
and

$$y^{*}_{RCA,i,t} = \frac{-\left(E^{*}_{RCA,E,i,t}b_{3,RCA,E} + E^{*}_{RCA,D,i,t}b_{3,RCA,D}\right)}{E^{*}_{RCA,E,i,t}b_{4,RCA,E} + E^{*}_{RCA,D,i,t}b_{4,RCA,D}}.$$

In Table II-6 $b_{3,RCA,E} > 0$, $b_{3,RCA,D} > 0$ and $b_{4,RCA,E} < 0$ but $b_{4,RCA,D} > 0$. Thus $y^{*}_{RCA,i,t}$ maximizes demand only if $E^{*}_{RCA,E,i,t}b_{4,RCA,E} + E^{*}_{RCA,D,i,t}b_{4,RCA,D} < 0$, i.e. if the share of electricity demand in the RCA sector exceeds 12%.

The simulated demand paths for a hypothetical country in Figure II-13 can be used to compare demand by sector. Figure II-14 plots total energy consumption in the industrial, RCA and transportation sectors. The solid line represents the long run equilibrium demand path where real GDP per capita grows at a constant rate of 2.15% per year. Figure II-14 illustrates that the pattern of economic development and energy demand in our hypothetical country is consistent with the theories of saturation and dematerialization, particularly in the transportation and industrial sectors. Although the share of demand is initially largest in the RCA sector, demand in the industrial sector grows rapidly during early stages of economic development. Because demand growth is steady but fairly low in the RCA sector, the share of industrial sector demand quickly becomes the largest share of total energy demand at approximately \$1,100 in our hypothetical country. However, by roughly \$38,000 industrial sector demand peaks and growth of the transportation sector propels it to be the largest share of total energy demand. Moreover, transportation sector demand growth continues until demand is maximized near \$135,522 which is beyond the income range in our sample.

Figure II-14: Simulated demand by sector for a hypothetical country



As mentioned previously, industrial demand peaks at an income level, $y_{Ind}^{*\max}$, near $38,000. Because electricity demand comprises only a small share of industrial electricity consumption (approximately 22%) in our hypothetical country, $y_{Ind}^{*\max}$ is closer to the high end of the range ($y_{Ind,DU}^{*\max}$ = $40,444) than the lower end of the range ($y_{IndE}^{*\max} = \$34,032$). In contrast, total energy demand in the RCA sector does not peak for our hypothetical country over any relevant GDP per capita range. Although electricity demand in the sector peaks at an income of $50,067 in Figure II-13, the share of electricity demand is small. The growth of RCA direct use fuel demand dominates any electricity demand loss beyond $50,067, causing demand in the RCA sector to continue to grow.

Allowing real GDP growth to vary annually (with mean of 2.15% and standard deviation of 3.38%) results in the demand path represented by hollow circles in Figure II-13. The deviation from the long run equilibrium demand path illustrates the impact of changing capital utilization in each sector. Figure II-14 illustrates that the impacts of capital utilization are modest in the RCA sector. However, significant deviations in the short run from the long run equilibrium are present in both the transportation and industrial sectors. This is consistent with increasing capital utilization causing short-term increases in energy intensity. Assuming both labor and energy are variable inputs, then less efficient capital stock and additional employees will be used to increase output, resulting in more transportation of the workforce and finished goods.

Comparison of Electricity and Direct-Use Fuel Demand

Equation (II-5) can also be used to generate long run models of electricity and direct-use fuel demand, as well as a model of total energy consumption. Electricity demand is the sum of electricity consumption in the industrial and RCA sectors.

$$\begin{aligned} E^*_{E,i,t} &= E^*_{Ind,E,i,t} + E^*_{RCA,E,i,t} \\ &= A_{Ind,E,i,t}\, p_{Ind,E,i,t}^{b_{1,Ind,E}}\, pop_{i,t}^{b_{2,Ind,E}}\, Y_{i,t}^{b_{3,Ind,E}} \exp(y_{i,t})^{b_{4,Ind,E}} \exp(hdd_{i,t}^2)^{b_{5,Ind,E}} \\ &\quad + A_{RCA,E,i,t}\, p_{RCA,E,i,t}^{b_{1,RCA,E}}\, pop_{i,t}^{b_{2,RCA,E}}\, Y_{i,t}^{b_{3,RCA,E}} \exp(y_{i,t})^{b_{4,RCA,E}} \exp(hdd_{i,t}^2)^{b_{5,RCA,E}} \end{aligned} \tag{II-21}$$

Direct-use fuel demand is the sum of transportation fuel and industrial and RCA direct-use fuel consumption.

$$\begin{aligned} E^*_{D,i,t} &= E^*_{Ind,D,i,t} + E^*_{RCA,D,i,t} + E^*_{Tran,i,t} \\ &= A_{Ind,D,i,t}\, p_{Ind,D,i,t}^{b_{1,Ind,D}}\, pop_{i,t}^{b_{2,Ind,D}}\, Y_{i,t}^{b_{3,Ind,D}} \exp(y_{i,t})^{b_{4,Ind,D}} \exp(hdd_{i,t}^2)^{b_{5,Ind,D}} \\ &\quad + A_{RCA,D,i,t}\, p_{RCA,D,i,t}^{b_{1,RCA,D}}\, pop_{i,t}^{b_{2,RCA,D}}\, Y_{i,t}^{b_{3,RCA,D}} \exp(y_{i,t})^{b_{4,RCA,D}} \exp(hdd_{i,t}^2)^{b_{5,RCA,D}} \\ &\quad + A_{Tran,i,t}\, p_{Tran,i,t}^{b_{1,Tran}}\, pop_{i,t}^{b_{2,Tran}}\, Y_{i,t}^{b_{3,Tran}} \exp(y_{i,t})^{b_{4,Tran}} \exp(hdd_{i,t}^2)^{b_{5,Tran}} \end{aligned} \tag{II-22}$$

Finally, total energy demand is the sum of electricity and direct-use fuel in each of the three sectors.

$$E^*_{i,t} = E^*_{Ind,E,i,t} + E^*_{RCA,E,i,t} + E^*_{Ind,D,i,t} + E^*_{RCA,D,i,t} + E^*_{Tran,i,t}\,. \tag{II-23}$$

The long run elasticity of electricity demand is equal to the share-weighted elasticity of electricity demand by sector.

$$\varepsilon^{*Y}_{E,i,t} = \frac{E^*_{Ind,E,i,t}\varepsilon^{*Y}_{Ind,E,i,t} + E^*_{RCA,E,i,t}\varepsilon^{*Y}_{RCA,E,i,t}}{E^*_{E,i,t}}\,.$$

The income level at which the long run elasticity of electricity demand equals zero is

$$y^*_{E,i,t} = \frac{-\left(E^*_{Ind,E,i,t}b_{3,Ind,E} + E^*_{RCA,E,i,t}b_{3,RCA,E}\right)}{E^*_{Ind,E,i,t}b_{4,Ind,E} + E^*_{RCA,E,i,t}b_{4,RCA,E}}\,.$$

Because $b_{4,Ind,E} > 0$, $b_{3,RCA,E} > 0$, $b_{4,Ind,E} < 0$ and $b_{4,RCA,E} < 0$ in Table II-6, $y^*_{E,i,t}$ maximizes the long run equilibrium demand for energy in the industrial sector. Moreover, $y^{*\max}_{E,i,t} \geq y^{*\max}_{Ind,E}$ and $y^{*\max}_{E,i,t} \leq y^{*\max}_{RCA,E}$. Thus, industrial demand is maximized between \$34,038 and \$50,067, and the higher the share of electricity demand by the industrial sector, the lower $y^{*\max}_{E,i,t}$ falls within this range.

Similarly, the long run income elasticity of direct-use fuel demand is equal to

$$\varepsilon^{*Y}_{D,i,t} = \frac{E^*_{Ind,D,i,t}\varepsilon^{*Y}_{Ind,D} + E^*_{RCA,D,i,t}\varepsilon^{*Y}_{RCA,D} + E^*_{Tran,i,t}\varepsilon^{*Y}_{Tran}}{E^*_{i,t}}$$

Also,

$$y^*_{D,i,t} = \frac{-\left(E^*_{Ind,D,i,t}b_{3,Ind,D} + E^*_{RCA,D,i,t}b_{3,RCA,D} + E^*_{Tran,i,t}b_{3,Tran}\right)}{E^*_{Ind,D,i,t}b_{4,Ind,D} + E^*_{RCA,D,i,t}b_{4,RCA,D} + E^*_{Tran,i,t}b_{4,Tran}}.$$

In Table II-6 $b_{3,Ind,D} > 0$, $b_{3,RCA,D} > 0$, $b_{3,Tran} > 0$, $b_{4,Ind,D} < 0$ and $b_{4,Tran} < 0$ but $b_{4,RCA,D} > 0$. Thus $y^*_{D,i,t}$ maximizes demand only if $E^*_{Ind,D,i,t}b_{4,Ind,D} + E^*_{RCA,D,i,t}b_{4,RCA,D} + E^*_{Tran,i,t}b_{4,Tran} < 0$.

Finally, the long run income elasticity of total energy demand is equal to

$$\varepsilon^{*Y}_{D,i,t} = \frac{E^*_{Ind,D,i,t}\varepsilon^{*Y}_{Ind,D} + E^*_{Ind,E,i,t}\varepsilon^{*Y}_{Ind,E} + E^*_{RCA,E,i,t}\varepsilon^{*Y}_{RCA,E} + E^*_{RCA,D,i,t}\varepsilon^{*Y}_{RCA,D} + E^*_{Tran,i,t}\varepsilon^{*Y}_{Tran}}{E^*_{i,t}},$$

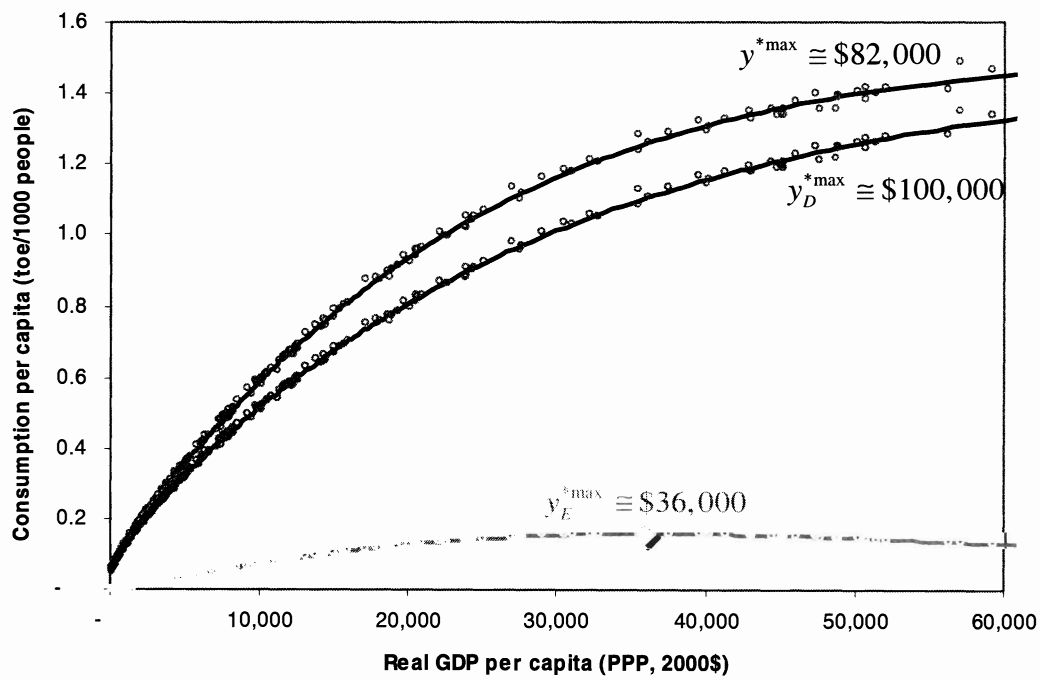
and

$$y^*_{D,i,t} = \frac{-\left(E^*_{Ind,E,i,t}b_{3,Ind,E} + E^*_{Ind,D,i,t}b_{3,Ind,D} + E^*_{RCA,E,i,t}b_{3,RCA,E} + E^*_{RCA,D,i,t}b_{3,RCA,D} + E^*_{Tran,i,t}b_{3,Tran}\right)}{E^*_{Ind,E,i,t}b_{4,Ind,E} + E^*_{Ind,D,i,t}b_{4,Ind,D} + E^*_{RCA,E,i,t}b_{4,RCE,D} + E^*_{RCA,D,i,t}b_{4,RCA,D} + E^*_{Tran,i,t}b_{4,Tran}}$$

is the income level which equates the long run elasticity of energy demand to zero. $y^*_{D,i,t}$ maximizes total energy demand only if the denominator is less than zero.

The simulated demand paths for a hypothetical country presented in Figure II-13 are used to compare demand by commodity and to analyze total energy demand. Figure II-15 plots total energy consumption in the industrial, RCA and transportation sectors. The solid line represents the long run equilibrium demand path where real GDP per capita grows at a constant rate of 2.15% per year. Allowing real GDP growth to vary annually (with mean of 2.15% and standard deviation of 3.38%) results in the demand path represented by hollow circles. The deviation from the long run equilibrium demand path illustrates the impact of changing capital utilization in each sector.

Figure II-15: Simulated energy demand by commodity for a hypothetical country



In Figure II-15 , the simulated paths of electricity and direct-use fuel demand are each consistent with the theories of saturation and dematerialization. Electricity demand grows rapidly during early states of economic development. However, once income per capita reaches approximately $36,000, demand for electricity in our hypothetical country begins to fall. Direct-use fuel consumption also exhibits a non-monotonic relationship with economic development. Recall that direct-use demand is maximized in the industrial sector at $50,067 and $135,522 in the transportation sector. However, RCA direct-use fuel demand continues to grow with economic development over any relevant income range. The decrease in industrial direct-use fuel consumption eventually outweighs demand growth in both the RCA and transportation sectors, causing direct-use fuel demand in our hypothetical country to be maximized at approximately $100,000 per capita.[52] This leads us to conclude that the impacts of dematerialization and saturation are stronger for electricity than direct-use fuel, causing electricity demand to be maximized at a significantly lower income level than direct-use fuel. Moreover, because the share of direct-use fuel in our hypothetical country dwarfs that of electricity demand, total energy demand is not maximized until an income level of approximately $82,000.

Although direct-use fuel takes a greater share of demand than electricity in both the RCA and industrial sector during all stages of development in our hypothetical country, we find that the long run elasticity of demand, ε_E^{*Y}, to be initially greater than the income elasticity of demand for direct-use fuels, ε_D^{*Y}, i.e. $\varepsilon_E^{*Y} < \varepsilon_D^{*Y}$. This suggests that as an economy begins to develop, significant investments are made in electricity transmission and distribution to capture from economics of scale created from using a large generator to convert direct-use fuel into electricity. This allows electricity demand in the industrial and RCA sectors to grow rapidly. However, once income reaches approximately $20,000 and transportation becomes the largest share of demand by

[52] Although energy demand in the transportation sector peaks at $135,522 per capita in our hypothetical country, income elasticity in the sector is still positive around $100,000. Thus the fall in industrial direct-use demand must offset the growth in demand from both the RCA and transportation sectors near $100,000.

commodity-sector, the long run income elasticity of demand for direct-use fuel becomes greater than the income elasticity of electricity, i.e. $\varepsilon_E^{*Y} < \varepsilon_D^{*Y}$. This is consistent with differing physical characteristics and the current state of technology which prevent electricity from being a perfect substitute for direct-use fuel, particularly in the transportation sector.

j. Conclusions

In modeling energy consumption by commodity and sector for 50 countries, we find the relationship between energy consumption and economic development corresponds to the structure of aggregate output and the nature of derived demand for electricity and direct-use fuels in each sector. Notably, the evidence of non-constant income elasticity of demand is much greater for electricity demand than for direct-use fuel consumption. In addition, we show that during periods of rapid economic development, one in which the short-term growth rate exceeds the long-run average, an increase in aggregate output is met by less energy-efficient capital. This is a result of capital being fixed in the short-term. As additional, more efficient capital stock is added to the production process, the short-term increase in energy intensity will diminish.

k. References

Adkins, L.C. and R.C. Hill (2008). *Using Stata for Principles of Econometrics*. 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc.

Ahn, S. and P. Schmidt (1995). “Efficient Estimation of Models for Dynamic Panel Data.” *Journal of Econometrics*, 68(1): 5-27.

Anderson, T. and C. Hsiao (1981). “Estimation of Dynamic Models with Error Components.” *Journal of American Statistical Associations*, 76: 598-606.

Anderson, T. and C. Hsiao (1982). "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics*, 18: 67-82.

Ang, B.W. (1987). "A Cross-Sectional Analysis of Energy-Output Correlation." *Energy Economics*, 9 (4): 274-286.

Ang, B.W. (1988). "Electricity-Output Ratio and Sectoral Electricity Use." *Energy Policy*, 16(2): 115-121.

Arellano, M (1987). "Computing Robust Standard Errors for Within-Groups Estimators." *Oxford Bulletin of Economics & Statistics*, 49(4): 431-434.

Arellano, M. and S. Bond (1991). "Some Tests of Specification for Panel Data: Monte Carol Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2), 277-297.

Arellano, M. and O. Bover (1995). "Another Look at Instrumental Variables Estimation of Error Components Models." *Journal of Econometrics*, 68: 29-51.

Balestra, P. and M. Nerlove (1966). "Pooling Cross Setction and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas." *Econometrica*, 34(3): 585-612.

Baltagi, B.H. (2008). *Econometric Analysis of Panel Data.* 4th ed. Chichester, UK: John Wiley & Sons Ltd.

Baltagi, B.H. and J.M. Griffin (1997). "Pooled Estimators vs. Their Heterogeneous Counterparts in the Context of Dynamic Demand for Gasoline." *Journal of Econometrics,* 77: 303-327.

Baltagi, B and Q. Li (1991). "A Joint Hypothesis Test for Serial Correlation for Serial Correlation and Random Individual Effects." *Statistics & Probability Letter*, 11: 277-280.

Bernardini, O. and R. Galli (1993). "Dematerialization: Long-Term Trends in the Intensity of Use of Materials and Energy." *Futures*, (May): 431-448.

Breusch, T.S. and A.R. Pagan (1980). "The Lagrange Multiplier Test and its Applications to Model Specification in the Presence of Local Misspecification." *Review of Economic Studies*, 47: 239-253.

Brookes, L.G. (1973). "More on the Output Elasticity of Energy Consumption." *Journal of Industrial Economics*, (April): 83-94.

Blundell, R. and S. Bond (1988). "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics,*84(1): 115-143.

Chenery, H. B. and M. Syrquin (1975). *Patterns of Development, 1950-1970.* London: Oxford Press.

Darrat, A.F., O.W. Gilley, and D.J. Meyer (1996). "US Oil Consumption, Oil Prices, and the Macroeconomy." *Empirical Economics*, 21: 317-334.

Ferguson, R., W. Wilkinson and R. Hill (2000). "Electricity Use and Economic Development." *Energy Policy* 28: 923-934.

Galli, R. (1998). "The Relationship between Energy Intensity and Income Levels: Forecasting Long Term Energy Demand in Asian Emerging Countries." *The Energy Journal*, (19)4: 85-105.

Green, W.H. (2003). *Econometric Analysis.* 5th ed. New Jersey: Prentice-Hall.

Hankinson, G.A. and J.H.W. Rhys (1983). "Electricity Consumption, Electricity Intensity and Industrial Structure." *Energy Economics*, 5 (3): 146-152.

Hsiao, C. (1986). *Analysis of Panel Data.* Cambridge University Press.

Judson, R.A., R. Schmalensee and T.M. Stoker (1999). "Economic Development and the Structure of the Demand for Commercial Energy." *The Energy Journal*, (20)2: 29-57.

Koyck, L.M. (1954). *Distributed Lags and Investment Analysis.* Amsterdam: North Holland Publishing.

Kuznets, S. (1971). *Economic Growth of Nations: Total Output and Production Structure.* Cambridge, MA: The Belnapp Press of Harvard University.

Medlock III, K.B. and R. Soligo (2001). "Economic Development and End-Use Energy Demand." *The Energy Journal*, (22)2: 77-105.

Moroney, J.R. (1992). "Energy, Capital, and Growth" in *Advances in the Economics of Energy and Resources*, Volume 7: 189-204.

Nilsson, L.J. (1993). "Energy Intensity Trends in 31 Industrial and Developing Countries 1950-1988." *Energy*, 18 (4): 309-322.

StataCorp (2007). *Stata Longitudinal/Panel-Data Reference Manual: Release 10.* College Station, TX: StataCorp LP.

StataCorp (2007). *Stata Statistical Software: Release 9.* College Station, TX: StataCorp LP.

Stern, D. (1993). "Energy Use and Economic Growth in the USA, A Multivariate Approach." *Energy Economics*, 15: 137-150.

Stern, D (2000). "A Multivariate Cointegration Analysis on the Role of Energy in the US Macroeconomy." *Energy Economic*s, 15: 267-283.

Wolde-Rufael, Y. (2006). "Electricity Consumption and Economic Growth: a Time Series Experience for 17 African Countries." *Energy Policy*: 34: 1006-1114.

Wooldride, J. (1990). "A Note on the Lagrange Multiplier and F-Statistics for Two Stage Least Squares Regressions." *Economics Letters*, 34: 151-155.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: The MIT Press.

Zilberfarb, B.Z. and F.G. Adams (1981). "The Energy-GDP Relationship in Developing Countries." *Energy Economics*, 3: 244-248.

III. Essay 3 – Florida Vehicle Fuel Demand Decomposed into Vehicle Miles Traveled and Fuel Efficiency

a. Introduction

In this paper, we develop a system of equations to estimate a model of motor vehicle fuel consumption, vehicle miles traveled and implied fuel efficiency for the 67 counties of the State of Florida from 2001 to 2008. This procedure allows us to decompose the factors of fuel demand into elasticities of vehicle driving demand and fuel efficiency. Particular attention is paid to the influence of the price of fuel, the sale of goods and services, vehicle ownership and population density on each component of our model.

In summary, we find that an increase in the price of fuel results in a short and long-run decrease in fuel demand but not through a decrease in vehicle miles traveled. Instead we find evidence that a price increase leads to an increase in fuel efficiency. We also find that the value of goods sold is positively correlated with vehicle fuel demand as consumers likely increase vehicle miles traveled to acquire more goods. However, this has no significant impact on vehicle fuel efficiency. In addition, we find that an increase in the share of vehicle ownership does not result in an increase in vehicle miles traveled in our sample. Alternatively, a decrease in vehicle fuel efficiency causes vehicle fuel demand to increase. Finally, an increase in population density decreases vehicle miles traveled but the vehicle fuel savings is offset by 65% in the short-run, and 35% in the long-run, by a loss of fuel efficiency.

b. Determinants of Vehicle Fuel Demand

Refined petroleum products, gasoline and diesel fuel, are the main energy inputs into the production of personal and commercial transportation services.[53] Generally, a household or individual will maximize utility gained from consuming goods and services where transportation is an inputs into the household production function. [54] The production function therefore determines the indirect demand for fuel subject to the wage rate, time spent driving, the technology used to convert fuel into distance traveled and the price of fuel and all other goods and services. Because fuel demand (measured in gallons) is simply the outcome of vehicle miles traveled divided by vehicle fuel efficiency (measured in miles per gallon), it is possible to decompose vehicle fuel demand into the demand for vehicle miles traveled and fuel efficiency. This form of fuel demand model is described in detail in section c. First, however, we identify several factors that are likely to influence Florida residents' driving demand and fuel efficiency.

Price, Economic Activity and Vehicle Ownership

As alluded to above, it is standard to assume gasoline consumption is a function of income and price; and studies generally find estimates of price elasticity to be negative

[53] According to the EIA, the use of alternative fuels such as compressed natural gas and electricity accounted for less than 0.23% of total vehicle fuel consumed in the United States in 2007. Gasoline and diesel (and additives) account for the remaining vehicle fuel consumed. Note that a single engine generally cannot use multiple fuel types. Thus, gasoline and diesel are substitutes only in the long-run. Also, one gallon of diesel has a higher heat content (approximately 139,000 btu/gallon) than gasoline (approximately 124,000 btu/gallon). Unless otherwise stated, all volumes have been converted to gallon of gasoline equivalent where 1 gallon of diesel is equal to 1.12 gallons of gasoline.

[54] A thorough review of this utility framework can be found in Sterner and Dahl (1992). Other examples include Puller and Greening (1999) and Mehta et al (1978). An alternative specification would include an individual maximizing utility gained from consuming goods and transportation subject to income and prices of vehicle fuel and all other goods and services.

and income elasticity to be positive.[55] In a review of 97 studies, Dahl and Sterner (1991) report the short and long-run price elasticity estimates average -0.26 and -0.86, respectively, whereas the short and long-run income elasticity estimates average 0.48 and 1.38, respectively. Epsey (1998) finds similar results in a review of 101 gasoline demand studies. Specifically, the 363 short-run and 277 long-run price elasticity estimates average -0.26 and -0.58, respectively. The 345 short-run and 245 long-run income elasticity estimates average 0.47 and 0.88, respectively. Espey (1998) also concludes that the elasticity estimates are sensitive to the inclusion or exclusion of vehicle ownership and fuel efficiency. In general, excluding these variables produces more elastic estimates for price and income.

In the context of a fuel demand model decomposed into vehicle miles traveled and fuel efficiency, a driver can respond to a change in the price of fuel by altering vehicle miles traveled. This is a result of the fact that vehicle fuel is an input and represents a cost in the consumer's transportation decision. For example, if the price of gasoline increases, an individual wanting to reduce gasoline expenses can substitute the use of his personal vehicle for public transportation (if available), carpool to work, reduce distance traveled by aggregating trips to several destinations into one (although this may reduce fuel efficiency) or eliminating less valuable travel such as driving vacations. Notice that these are predominately short-run responses. In the long-run he may move closer to work or change jobs to reduce the length of his commute for example.

Vehicle fuel efficiency – the conversion of gallons of fuel into vehicle miles traveled – is not only a function of vehicle characteristics (engine technology, size,

---

[55] Models excluding price and income variables are considered mis-specified by Dahl and Sterner (1991), for example.

weight, etc.) but also a function of vehicle maintenance (tire pressure, regular oil changes, etc.) as well as driving conditions (traffic, speed, terrain, etc.). Thus, in the short-run where vehicle stocks are assumed to be fixed, a driver can reduce gasoline expenses by increasing the frequency of routine maintenance or altering his route to avoid roadway congestion and stop-and-go traffic. A driver may also reduce his driving speed to optimize engine performance, although this has the added cost of increasing driving time. Alternatively, in the case of a multi-vehicle household, the members may chose to reallocate the use of vehicles among themselves to minimize total fuel expenses. Finally, in the long-run, an individual can purchase a new vehicle with improved engine technology, reduce consumption of motor vehicle fuel and maintain the same quantity vehicle miles traveled. Note, however, there may be additional costs such as reduced comfort or safety.

Using household survey data from the 1980s, Puller and Greening (1999) find the price elasticity of both gasoline consumption and vehicle miles to be negative. This is of no surprise. However, the authors also find the price elasticity of fuel efficiency (measured in miles per gallon) to be negative and suggest that households respond to higher gasoline prices by reducing high efficiency miles (such as the family vacation) rather than improving its vehicle stock or altering driving behavior. The same result may also be achieved if the household decides to combine several trips into one, resulting in fewer miles traveled. However, additional stops may reduce overall fuel efficiency.

Income is thought to impact vehicle fuel demand in several ways. First, higher income is associated with a higher rate of vehicle ownership, a necessity for personal vehicle transportation.[56] Thus, following Espey (1998), we explicitly include vehicle ownership in our model of vehicle fuel demand, vehicle miles traveled and implied fuel efficiency. A positive elasticity of demand for vehicle miles traveled will suggest, provided the household owns a vehicle, more demand for miles driven. However, a contrary effect might operate if vehicle ownership decreases fuel efficiency, for example because an increase in the number of vehicles on the road increases congestion and results in poorer engine performance. Another way this might happen would be if safety or roominess of the vehicle, for example, have positive income elasticities and vehicle safety and size are negatively correlated with vehicle fuel efficiency. Wage rates, and thus the opportunity cost of time, might provide another link between income and fuel efficiency. Individuals with a higher cost of time might substitute fuel for time by driving faster and thus purchase vehicles that can go faster even though they may be less fuel-efficient.

Another issue is how one measures income at the aggregate level. Data on residents' median household income is available at the county level. Holding all else equal, an increase in household income implies an individual can afford to purchase more vehicle fuel. In the case of our study, however, Florida has a large retirement community and the use of household income data may not accurately capture fuel expenditures out of wealth. Alternatively, GDP per capita is often used. Here an increase in GDP per capita signals both an increase in income and an increase in economic activity, which could result in an increase in vehicle miles traveled as consumers purchase more goods and

[56] Examples of empirical work include Medlock and Soligo (2002) and Dargay, Gately and Sommer (2007).

services.[57] Furthermore, the use of GDP per capita at the county level would likely capture demand for transportation services by out-of-county residents such as tourists and commuters who spend money while in the county despite being a resident elsewhere. This is unlikely to be captured by using the median household income of all residents of the county. Unfortunately, GDP per capita data at the county level is not generally available. Instead we use county-level sales data, which is available by county in the State of Florida.

Population Density

Higher population density is likely to reduce demand for personal vehicle transportation. First, Stewart and Bennett (1975) suggest that residents of urban, more densely populated areas generally have greater access to public transportation, providing a substitute for personal vehicle transportation.[58] Second, the distance an individual must travel when commuting to work, or running routine errands such as purchasing groceries, may be lower in larger urban areas with multiple shopping and employment centers.[59]

Using 2008 Florida county-level data, Figure III-1 plots the natural logarithm of population density against the natural logarithm of average daily vehicle miles traveled per resident.[60] A visual inspection of the data reveals a strong negative correlation between population density and vehicle miles traveled. This is consistent with a review of the literature.

---

[57] For cross-country or long time-series data, higher per capita GDP or higher median household income may also be correlated with increased urbanization and access to better road conditions and infrastructure such as gas stations. Presumably, this further supports the correlation of income and driving demand.
[58] Also see Banister (1992) and Newman and Kenworthy (1989).
[59] Also see Ostro and Naroff (1980), Zelinsky and Sly (1984), and Bento et al (2005).
[60] The data is described in detail in section d. Varying marker colors are for illustration purposes only. The 33 counties that do not comprise one of the 20 metropolitan statistical areas (MSAs) in the state are labeled in blue. Counties labeled in green have the highest population in the MSA, and all remaining counties are labeled in red.

Figure III-1: Florida Daily Vehicle Miles Traveled per Resident and Density, 2008

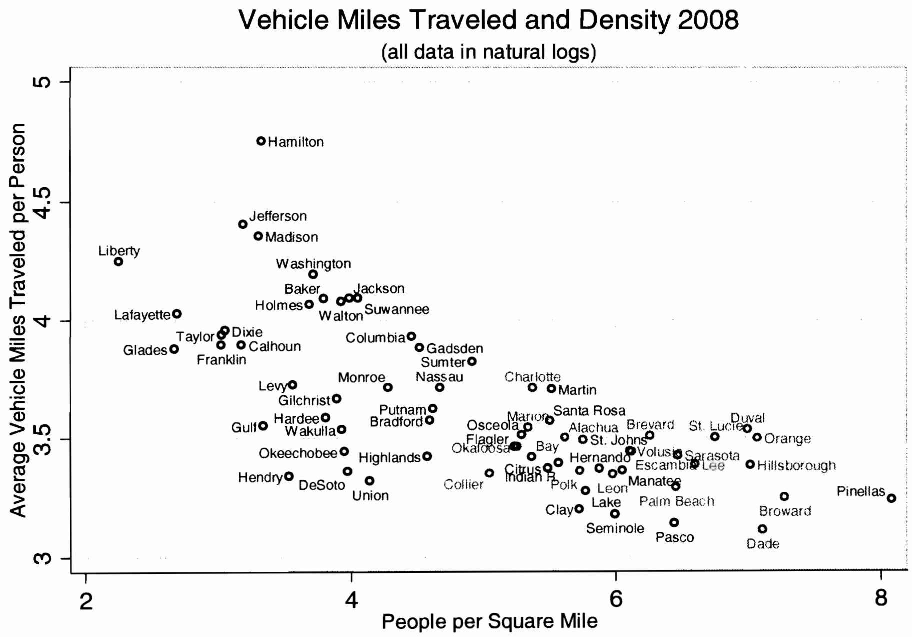


Figure III-2: Florida Implied Fuel Efficiency and Density, 2008

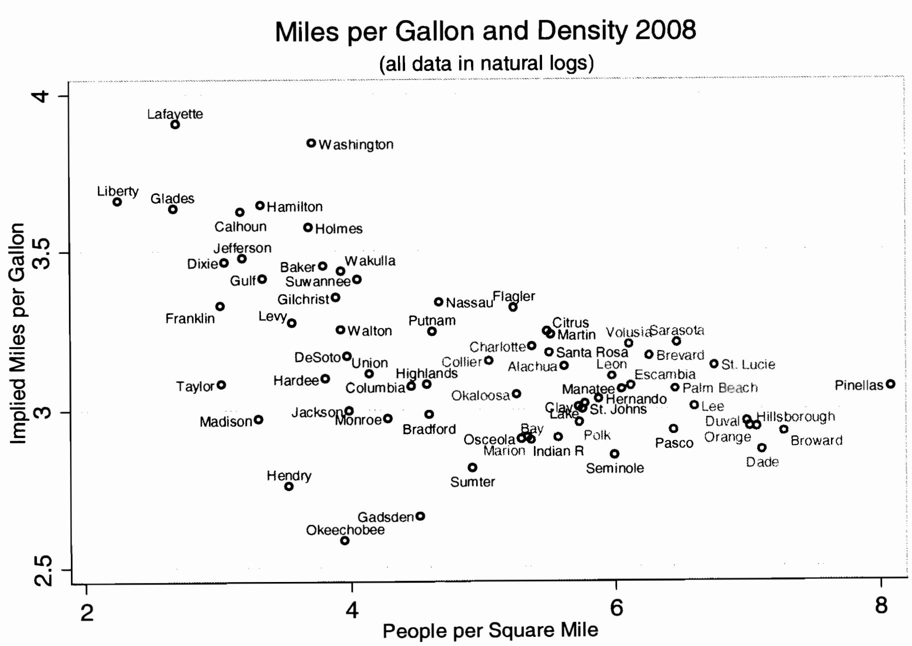
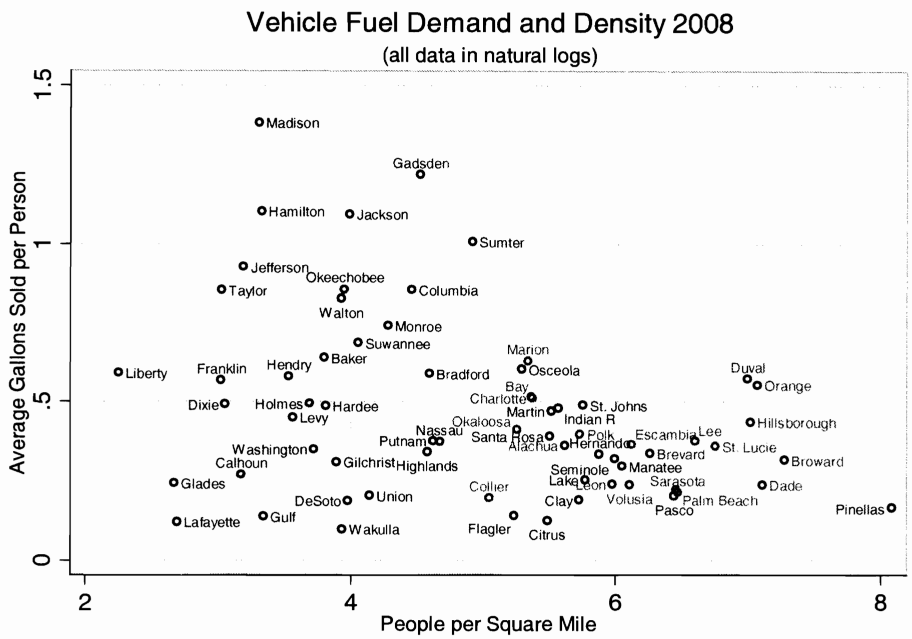


Figure III-3: Florida Daily Fuel Demand per Resident and Population Density, 2008



Stewart and Bennett also argue that road congestion and traffic tend to reduce fuel efficiency as density increases.[61] In Figure III-2, the natural logarithm of population density is plotted against the natural logarithm of implied vehicle fuel efficiency measured in miles per gallon. A visual inspection of data reveals there is a negative correlation between Florida population density and implied fuel efficiency.

Therefore, in terms of the quantity of vehicle fuel consumed, the literature as well as the data suggests that the decrease in personal vehicle transportation as population density increases is at least partially offset by a loss of fuel efficiency. Empirical results from Sanghi (1976), Houthakker, Verleger, and Sheehan (1976) and Newman and Kenworthy (1989) indicate the offset is less than complete so that gasoline demand nevertheless decreases with population density. This conclusion is less straightforward,

---

[61] Also see Banister (1992).

however, in our analysis of vehicle fuel sales in Florida. See Figure II-8 which plots the natural logarithm of population density against the natural logarithm of the average number of gallons of fuel sold per county resident. In 2008 the county with the highest and lowest motor fuel sales per capita (Madison and Wakulla, respectively) are generally lower density counties.

Although the literature on U.S. gasoline demand is quite mature, the majority of this research is focused on household, state or national-level consumption.[62] There are a few exceptions, however, particularly from authors interested in analyzing the affect of population density on gasoline consumption using metropolitan statistical area (MSA) data.[63] Unfortunately, these studies exclude rural areas, which are an important segment of the distribution of population density. Furthermore, county-level gasoline demand data – which would include rural areas – is rare.[64] A survey of available data by Zellinsky and Sly (1984) led the authors to believe that "no source, or combination of sources, provides a reasonable approximation of the amount of gasoline used for all purposes, and for personal transportation in particular, in the United States, by county, for 1960 and 1970."

By focusing on the 67 counties in the State of Florida, we are able to have a large sample size with a wide range of population densities. In 2008, population density in our sample spans from only 9.5 people per square mile in Liberty County to 3,250.9 people per square mile in Pinellas County. See

Table III-1 for a list of Florida counties, population, land area and density.

[62] See Table 2 and Table 3 in Espey (1998).
[63] Stewart and Bennett (1975) and Houthakker, Verlerger and Sheehan (1974), for example.
[64] In general we found states which grant counties the authority to tax gasoline are more likely to collect county-level gasoline sales data. According to the U.S. Energy Information Administration (2009) only Hawaii, Nevada and Florida grant this authority.

Table III-1: Florida counties, population, area and density

County	Population	Land Area	Density	County	Population	Land Area	Density
	count	*sq. miles*	*people per sq. mile*		*count*	*sq. miles*	*people per sq. mile*
Liberty	7,957	836	9.5	Flagler	91,247	485	188.1
Glades	11,175	774	14.4	Okaloosa	179,693	936	192.0
Lafayette	8,013	543	14.8	Osceola	263,676	1,322	199.5
Franklin	11,202	544	20.6	Marion	329,628	1,579	208.8
Taylor	21,546	1,042	20.7	Bay	163,946	764	214.6
Dixie	14,957	704	21.2	Charlotte	150,060	694	216.2
Calhoun	13,617	567	24.0	Citrus	141,416	584	242.2
Jefferson	14,547	598	24.3	Santa Rosa	150,053	609	246.4
Madison	18,895	692	27.3	Martin	138,660	556	249.4
Hamilton	14,348	515	27.9	Indian River	132,315	503	263.1
Gulf	15,667	555	28.2	Alachua	241,364	874	276.2
Hendry	39,453	1,153	34.2	Clay	184,727	601	307.4
Levy	39,460	1,118	35.3	Polk	580,594	1,874	309.8
Holmes	19,328	483	40.0	St. Johns	181,540	572	317.4
Washington	23,928	580	41.3	Lake	307,243	953	322.4
Baker	26,164	585	44.7	Hernando	171,689	478	359.2
Hardee	28,888	637	45.4	Leon	264,063	667	395.9
Gilchrist	17,191	349	49.3	Seminole	410,854	1,017	404.0
Walton	53,837	1,058	50.9	Manatee	315,766	741	426.1
Wakulla	31,089	607	51.2	Volusia	498,036	1,103	451.5
Okeechobee	40,359	774	52.1	Escambia	302,939	662	457.6
DeSoto	33,991	637	53.4	Brevard	536,521	1,018	527.0
Jackson	49,656	916	54.2	Pasco	471,028	745	632.3
Suwannee	39,802	688	57.9	Palm Beach	1,265,293	1,974	641.0
Union	15,141	240	63.1	Sarasota	372,057	573	649.3
Monroe	72,243	997	72.5	Lee	593,136	804	737.7
Columbia	69,092	797	86.7	St. Lucie	265,108	308	860.7
Gadsden	47,560	516	92.2	Duval	850,962	774	1,099.4
Highlands	100,011	1,028	97.3	Hillsborough	1,180,784	1,051	1,123.5
Bradford	29,012	293	99.0	Orange	1,072,801	908	1,181.5
Putnam	73,459	722	101.7	Dade	2,398,245	1,946	1,232.4
Nassau	69,835	652	107.1	Broward	1,751,234	1,205	1,453.3
Sumter	74,721	546	136.9	Pinellas	910,260	280	3,250.9
Collier	315,258	2,025	155.7				

The data is described in detail in section d.

We are also able build upon the previous research by developing a system of equations to model motor vehicle fuel consumption per resident, vehicle miles traveled per resident and implied vehicle fuel efficiency in the 67 counties in the State of Florida. Because we model total vehicle miles traveled, our vehicle fuel demand data must incorporate both gasoline and diesel fuel consumption. With this system of equations we are able to decompose fuel demand elasticity estimates into elasticities of demand for vehicle miles traveled and motor vehicle fuel efficiency, and analyze the ways various factors influence fuel demand. Particular attention is paid to price of vehicle fuel, economic activity, vehicle ownership and population density.

c. Model of Fuel Demand, Vehicle Miles Traveled and Fuel Efficiency

In this section, a dynamic system of structural equations for annual vehicle fuel demand (measured in gallons), annual personal vehicle transportation demand (measured in vehicle miles traveled) and fuel efficiency (measured in miles per gallon) is described in detail.

Equation 1: Demand for Vehicle Fuel

Our first equation describes demand for vehicle fuel directly and assumes the long-run equilibrium demand per resident measured in gallons ($g^*_{i,t}$) is a function of the price of a gallon of fuel ($p_{i,t}$), gross sale of goods and services ($s_{i,t}$), vehicles per resident ($v_{i,t}$) and population density ($d_{i,t}$) where i and t denote county and year, respectively. This can be written as

$$g^*_{i,t} = f\left(p_{i,t}, s_{i,t}, v_{i,t}, d_{i,t}\right). \tag{III-1}$$

We use a the following demand function assuming constant elasticity.

$$g^*_{i,t} = A_{1,i} p_{i,t}^{b_{1,1}} s_{i,t}^{b_{1,2}} v_{i,t}^{b_{1,3}} d_{i,t}^{b_{1,4}} \tag{III-2}$$

Taking the natural logarithm of (III-2) yields

$$\ln g_{i,t}^{*} = c_1 + a_{1,i} + b_{1,1} \ln p_{i,t} + b_{1,2} \ln s_{i,t} + b_{1,3} \ln v_{i,t} + b_{1,4} \ln d_{i,t} \tag{III-3}$$

where $b_{1,k}$ for $k = 1,...,4$ and $c_1 + a_{1,i} = \ln A_{1,i}$ are parameters of the long-run equilibrium fuel demand equation. Note that c_1 is the common intercept where as $a_{1,i}$ is the county-specific effect, which can be treated as either fixed or random.

Responses to exogenous changes in our equation of fuel demand, however, are unlikely to occur in just one year.[65] To capture both short and long-run responses, we use the partial adjustment model specified by Houthakker, Verleger, and Sheehan (1974) where the adjustment to the long-run equilibrium is assumed to take the form

$$\left(\frac{g_{i,t}}{g_{i,t-1}} \right) = \left(\frac{g_{i,t}^{*}}{g_{t-1}} \right)^{\gamma_{1,1}} \tag{III-4}$$

and $\gamma_{1,1}$ is the adjustment factor which is bound between 0 and 1.

Taking the natural logarithm of (III-4) and rearranging the terms yields

$$\ln g_{i,t} = \gamma_{1,1} \ln g_{i,t}^{*} + \left(1 - \gamma_{1,1}\right) \ln g_{i,t-1}. \tag{III-5}$$

It is now obvious that if $\gamma_{1,1} = 0$, the adjustment to long-run equilibrium never happens while if $\gamma_{1,1} = 1$, the adjustment is instantaneous. Substituting (III-3) into (III-5), the first equation of our dynamic system describing fuel demand is

$$\begin{aligned} \ln g_{i,t} = \mu_1 + \alpha_{1,i} + \beta_{1,1} \ln p_{i,t} + \beta_{1,2} \ln s_{i,t} + \beta_{1,3} \ln v_{i,t} + \beta_{1,4} \ln d_{i,t} \\ + \left(1 - \gamma_{1,1}\right) \ln g_{i,t-1} + u_{1,i,t} \end{aligned} \tag{III-6}$$

[65] In a survey of over one hundred studies of the short and long-run elasticities of gasoline demand, Dahl and Sterner (1991) find that "static models tend to underestimate long-run adjustments to price changes but not necessarily to income changes."

where the short-run parameter estimates, $\beta_{1,k} = \gamma_{1,1} b_{1,k}$ for $k = 1,\ldots,4$, $\mu_1 = \gamma_{1,1} c_1$ and $\alpha_{1,i} = \gamma_{1,1} a_{1,i}$, can be used to recover the long-run parameter estimates, $a_{1,k}$, $b_{1,k}$ and c_1.

Equation 2: Demand for Vehicle Miles Traveled

The above approach is largely consistent with dynamic models of gasoline demand analyzed in surveys by Dahl and Sterner (1991) and Espey (1998). However, fuel demand can also be disaggregated into its components – vehicle miles traveled and fuel efficiency – to analyze elasticity estimates separately.[66]

Given fuel demand in equation 1 is implicitly derived from demand for vehicle transportation, we assume the long-run equilibrium demand for vehicle miles traveled per capita, $m_{i,t}$, takes a similar constant elasticity log-linear form. Specifically,

$$\ln m_{i,t}^{*} = c_2 + a_{2,i} + b_{2,1} \ln p_{i,t} + b_{2,2} \ln s_{i,t} + b_{2,3} \ln v_{i,t} + b_{2,4} \ln d_{i,t} . \tag{III-7}$$

The dynamic nature of vehicle transportation demand is captured by a partial adjustment process similar to (III-4), where

$$\left(\frac{m_{i,t}}{m_{i,t-1}} \right) = \left(\frac{m_{i,t}^{*}}{m_{t-1}} \right)^{\gamma_{2,2}} \tag{III-8}$$

and

$$\ln m_{i,t} = \gamma_{2,2} \ln m_{i,t}^{*} + \left(1 - \gamma_{2,2}\right) \ln m_{i,t-1} . \tag{III-9}$$

Substituting (III-7) into (III-9), the dynamic equation for vehicle miles traveled in our system is

[66] Examples include Dahl (1979), Archibald and Gillingham (1981) and Greening and Puller (1999). In Espey (1998), the author analyzes 277 short-run and 363 long-run models of gasoline demand of which 7 and 13, respectively, use an indirect approach where driving, vehicle ownership and fuel efficiency are analyzed separately.

$$\ln m_{i,t} = \mu_2 + \alpha_{2,i} + \beta_{2,1} \ln p_{i,t} + \beta_{2,2} \ln s_{i,t} + \beta_{2,3} \ln v_{i,t} + \beta_{2,4} \ln d_{i,t} + \left(1 - \gamma_{2,2}\right) \ln m_{i,t-1} + u_{2,i,t} \quad \text{(III-10)}$$

where $\beta_{2,k} = \gamma_{2,2} b_{2,k}$ for $k = 1,...,4$, the fixed or random effects are $\alpha_{2,i} = \gamma_{2,2} a_{2,i}$ and $\mu_2 = \gamma_{2,2} c_2$. This is the second structural equation of our dynamic system to be estimated.

Equation 3: Implied Fuel Efficiency

Because we now have equations for average vehicle fuel consumption per resident and average vehicle miles traveled per resident, we can derive an equation for the implied vehicle fuel efficiency measured in miles per gallon, $mpg_{i,t}$. By definition, it must be true that

$$mpg^*_{i,t} = \frac{m^*_{i,t}}{g^*_{i,t}} \quad \text{(III-11)}$$

in equilibrium. Taking the natural logarithm of (III-11) and substituting (III-7) for $\ln m^*_{i,t}$ and (III-3) for $\ln g^*_{i,t}$, the long-run equilibrium of implied vehicle fuel efficiency is

$$\ln mpg^*_{i,t} = c_3 + a_{3,i} + b_{3,1} \ln p_{i,t} + b_{3,2} \ln s_{i,t} + b_{3,3} \ln v_{i,t} + b_{3,4} \ln d_{i,t} \quad \text{(III-12)}$$

where $c_{3,i} = c_{2,i} - c_{1,i}$; $a_{3,i} = a_{2,i} - a_{1,i}$; $b_{3,k} = b_{2,k} - b_{1,k}$ for $k = 1,...,4$, $a_{3,i} = a_{2,i} - a_{1,i}$ and $b_{3,k} = b_{2,k} - b_{1,k}$ for $k = 1,...,4$.

To capture both long and short-run responses in fuel consumption and driving behavior, the fuel efficiency adjustment to the long-run equilibrium is assumed to take the following form.

$$\left(\frac{mpg_{i,t}}{mpg_{i,t-1}}\right)=\left(\frac{m^*_{i,t}}{m_{i,t-1}}\right)^{\gamma_{3,2}}\left(\frac{g^*_{i,t}}{g_{i,t-1}}\right)^{-\gamma_{3,1}} \tag{III-13}$$

Taking the natural logarithm of (III-13) and rearranging terms yields

$$\ln mpg_{i,t}=\gamma_{3,2}\ln m^*_{i,t}-\gamma_{3,1}\ln g^*_{i,t}+\left(1-\gamma_{3,2}\right)\ln m_{i,t-1}-\left(1-\gamma_{3,1}\right)\ln g_{i,t-1}. \tag{III-14}$$

Recall, however, the dynamic equations of fuel consumption and vehicle miles traveled. Substituting (III-5) and (III-9) into the identity describing fuel efficiency, $\ln mpg_{i,t}=\ln m_{i,t}-\ln g_{i,t}$, then it is also true that

$$\ln mpg_{i,t}=\gamma_{2,2}\ln m^*_{i,t}-\gamma_{1,1}\ln g^*_{i,t}+\left(1-\gamma_{2,2}\right)\ln m_{i,t-1}-\left(1-\gamma_{1,1}\right)\ln g_{i,t-1}. \tag{III-15}$$

From (III-14) and (III-15), it must be the case that $\gamma_{3,1}=\gamma_{1,1}$ and $\gamma_{3,2}=\gamma_{2,2}$ in our system of three equations.

Notice the form of adjustment in (III-13) differs from the fuel demand adjustment in (III-4) and the vehicle miles traveled adjustment in (III-8) the demand for vehicle miles traveled. Specifically, lagged vehicle miles traveled and lagged fuel demand are included in the vehicle fuel efficiency equation rather than lagged fuel efficiency. This is to allow the speed of adjustment to the long-run equilibrium to differ between fuel demand and vehicle miles traveled, i.e. the system does not require $\gamma_{1,1}=\gamma_{2,2}$. If, however, $\gamma_{1,1}=\gamma_{2,2}$, then $\gamma_{3,1}=\gamma_{3,2}$, and (III-14) reduces to

$$\ln mpg_{i,j}=\gamma_{3,3}\ln mpg^*_{i,t}+\left(1-\gamma_{3,3}\right)\ln mpg_{i,t-1} \tag{III-16}$$

where $\gamma_{3,1}=\gamma_{3,2}=\gamma_{3,3}$. Nonetheless, we allow the adjustment in each equation to vary and later test the restriction $\gamma_{3,1}=\gamma_{3,2}=\gamma_{3,3}$.

Finally, substituting the dynamic equations for fuel consumption and vehicle travel in (III-6) and (III-10), respectively, dynamic fuel efficiency can be described by the following equation.

$$\begin{aligned}\ln mpg_{i,t} = &\left(\mu_2 - \mu_1\right) + \left(\alpha_{2,i} - \alpha_{1,i}\right) + \left(\beta_{2,1} - \beta_{1,1}\right)\ln p_{i,t} + \left(\beta_{2,2} - \beta_{1,2}\right)\ln s_{i,t} \\ &+ \left(\beta_{2,3} - \beta_{1,3}\right)\ln v_{i,t} + \left(\beta_{2,4} - \beta_{1,4}\right)\ln d_{i,t} \\ &+ \left(1 - \gamma_{3,2}\right)\ln m_{i,t-1} - \left(1 - \gamma_{3,1}\right)\ln g_{i,t-1} + u_{3,i,t}\end{aligned} \tag{III-17}$$

Letting $\mu_3 = \mu_2 - \mu_1$, $\alpha_{3,i} = \alpha_{2,i} - \alpha_{1,i}$ and $\beta_{3,k} = \beta_{2,k} - \beta_{1,k}$ for $k = 1,\ldots,4$, the third and final equation that we wish to estimate is

$$\begin{aligned}\ln mpg_{i,t} = &\mu_3 + \alpha_{3,i} + \beta_{3,1}\ln p_{i,t} + \beta_{3,2}\ln s_{i,t} + \beta_{3,3}\ln v_{i,t} + \beta_{3,4}\ln d_{i,t} \\ &+ \left(1 - \gamma_{3,2}\right)\ln m_{i,t-1} - \left(1 - \gamma_{3,1}\right)\ln g_{i,t-1} + u_{3,i,t}\end{aligned}. \tag{III-18}$$

System of Equations Model

The system to be estimated is composed of three dynamic structural equations describing the demand for vehicle fuel, vehicle miles traveled and fuel efficiency. From (III-6), (III-10) and (III-18) the short-run system is

$$\begin{aligned}\ln g_{i,t} = &\mu_1 + \alpha_{1,i} + \beta_{1,1}\ln p_{i,t} + \beta_{1,2}\ln s_{i,t} + \beta_{1,3}\ln v_{i,t} + \beta_{1,4}\ln d_{i,t} \\ &+ \left(1 - \gamma_{1,1}\right)\ln g_{i,t-1} + u_{1,i,t} \\ \ln m_{i,t} = &\mu_2 + \alpha_{2,i} + \beta_{2,1}\ln p_{i,t} + \beta_{2,2}\ln s_{i,t} + \beta_{2,3}\ln v_{i,t} + \beta_{2,4}\ln d_{i,t} \\ &+ \left(1 - \gamma_{2,2}\right)\ln m_{i,t-1} + u_{2,i,t} \\ \ln mpg_{i,t} = &\mu_3 + \alpha_{3,i} + \beta_{3,1}\ln p_{i,t} + \beta_{3,2}\ln s_{i,t} + \beta_{3,3}\ln v_{i,t} + \beta_{3,4}\ln d_{i,t} \\ &+ \left(1 - \gamma_{3,2}\right)\ln m_{i,t-1} - \left(1 - \gamma_{3,1}\right)\ln g_{i,t-1} + u_{3,i,t}\end{aligned} \tag{III-19}$$

where several constraints follow from the identity $\ln mpg_{i,t} = \ln m_{i,t} - \ln g_{i,t}$.

$$\begin{aligned}&\text{Constraint A: } \alpha_{2,\mathrm{i}} - \alpha_{1,i} = \alpha_{3,i} \ \ \forall i = 1,\ldots,N\\&\text{Constraint B: } \mu_2 - \mu_1 = \mu_3\\&\text{Constraint C: } \beta_{2,k} - \beta_{1,k} = \beta_{3,k} \ \ \forall k = 1,\ldots,4\\&\text{Constraint D: } \gamma_{j,j} = \gamma_{3,j} \ \ \forall j = 1,2\end{aligned} \tag{III-20}$$

The parameter estimates of the long-run system of three equations

$$\begin{aligned}\ln g^*_{i,t} &= c_1 + a_{1,i} + b_{1,1}\ln p_{i,t} + b_{1,2}\ln s_{i,t} + b_{1,3}\ln v_{i,t} + b_{1,4}\ln d_{i,t}\\ \ln m^*_{i,t} &= c_2 + a_{2,i} + b_{2,1}\ln p_{i,t} + b_{2,2}\ln s_{i,t} + b_{2,3}\ln v_{i,t} + b_{2,4}\ln d_{i,t}\\ \ln mpg^*_{i,t} &= c_3 + a_{3,i} + b_{3,1}\ln p_{i,t} + b_{3,2}\ln s_{i,t} + b_{3,3}\ln v_{i,t} + b_{3,4}\ln d_{i,t}\end{aligned} \tag{III-21}$$

can be recovered from the estimation of the short-run system described by (III-19) and (III-20). In particular

$c_1 = \mu_1 / \gamma_{1,1}$, $c_2 = \mu_2 / \gamma_{2,2}$ and $c_3 = \mu_2 / \gamma_{3,2} - \mu_1 / \gamma_{3,1}$;

$a_{1,i} = \alpha_{1,i} / \gamma_{1,1}$, $a_{2,i} = \alpha_{2,i} / \gamma_{2,2}$ and $a_{3,i} = \alpha_{2,i} / \gamma_{3,2} - \alpha_{1,i} / \gamma_{3,1} \ \ \forall i = 1,\ldots,N$; and

$b_{1,k} = \beta_{1,k} / \gamma_{1,1}$, $b_{2,k} = \beta_{2,k} / \gamma_{2,2}$ and $b_{3,k} = \beta_{2,k} / \gamma_{3,2} - \beta_{1,k} / \gamma_{3,1} \ \ \forall k = 1,\ldots,4$.

Decomposition of the Elasticities of Fuel Demand

Using the parameter estimates of this system, fuel demand can be analyzed by component (demand for vehicle miles traveled and vehicle fuel efficiency) where

$$\ln g^*_{i,t} = \ln m^*_{i,t} - \ln mpg^*_{i,t}. \tag{III-22}$$

For example, consider any independent variable $x_{i,t}$ in our model of fuel demand. The long-run elasticity of fuel demand per capita with respect to variable $x_{i,t}$ (denoted $\varepsilon^{g^*}_x$), is equal to the elasticity of vehicle miles traveled (denoted $\varepsilon^{m^*}_x$), minus the elasticity of fuel efficiency (denoted $\varepsilon^{mpg^*}_x$). Formally,

$$\varepsilon_x^{g*} = \frac{\partial \ln g_{i,t}^*}{\partial x_{i,t}} \times x_{i,t} \tag{III-23}$$

where

$$\varepsilon_x^{m*} = \frac{\partial \ln m_{i,t}^*}{\partial x_{i,t}} \times x_{i,t}, \ \varepsilon_x^{mpg*} = \frac{\partial \ln mpg_{i,t}^*}{\partial x_{i,t}} \times x_{i,t} \tag{III-24}$$

and

$$\varepsilon_x^{g*} = \varepsilon_x^{m*} - \varepsilon_x^{mpg*}. \tag{III-25}$$

Therefore the long-run elasticity of fuel demand with respect to variable $x_{i,t}$ is positive if and only if $\varepsilon_x^{m*} > \varepsilon_x^{mpg*}$. Similarly, elasticity of fuel demand is negative if and only if $\varepsilon_x^{m*} < \varepsilon_x^{mpg*}$. Note the same holds in the short-run, where the short-run elasticity of fuel demand with respect to variable $x_{i,t}$ (denoted ε_x^{g}), is equal to the elasticity of vehicle miles traveled (denoted ε_x^{m}), minus the elasticity of fuel efficiency (denoted ε_x^{mpg}), i.e. $\varepsilon_x^{g} = \varepsilon_x^{m} - \varepsilon_x^{mpg}$.

d. Data

County-level data on motor gasoline and diesel consumption is made available by the Florida Energy and Climate Commission in its annual *Florida Motor Gasoline and Diesel Fuel Report*. The 2008 report contains monthly observations of motor gasoline sales "derived from local option gas tax receipts" beginning January 1989 through December 2008 for each of the 67 Florida counties.[67] These sales records, measured in gallons, were checked for accuracy and are consistent with U.S. Energy Information

[67] The local option gas tax is assessed in $/gallon and published by county in the *2008 Florida Motor Gasoline and Diesel*. Given the tax rate and county tax receipts, the calculation of gasoline sales measured in gallons is straightforward

Administration (2009) *State Energy Data System (SEDS)* statistics on motor gasoline consumed annually by the transportation sector in the State of Florida.[68]

Although annual diesel consumption measured in gallons is not available by calendar year at the county level, it is available by fiscal year (starting in July and ending in June). We analyzed monthly diesel fuel consumption at the state level and found first half diesel fuel consumption does not vary statistically from second half consumption for our sample. Thus, we assume county-level diesel fuel consumption for the calendar year is the simple average of the two relevant fiscal years. Finally, because the heat content of diesel fuel is approximately 139,000 btu/gallon whereas the heat content of gasoline is approximately 125,000 btu/gallon, diesel consumption is converted to a gallon of gasoline equivalent, i.e. 1 gallon of diesel is equal to 1.12 gallons of gasoline equivalent.[69]

Fuel demand in equation 1 is defined as the quantity of gasoline plus diesel consumed in each county measured in gallons of gasoline equivalent divided by the number of days per year and county population.[70] The system of structural equations is estimated using annual observations from 2001 to 2008.[71]

The Transportation Statistics Office at the Florida Department of Transportation collects and publishes an annual report on public roads. The report contains centerline miles (defined as the "length of road, without regard to the number of lanes") and daily vehicle miles traveled (defined as the "the product of a road's centerline miles and its

---

[68] The Energy Information Administration's *SEDS* contains state-level data only.
[69] While consumption data by fuel type is available, data on vehicle miles traveled is only provided at the aggregate level. Otherwise, one would be able to develop separate models of gasoline and diesel fuel demand.
[70] County population is available from the U.S. Census Bureau, Population Division (2009).
[71] The model is restricted to yearly observations from 2001 to 2008 due data limitations for several of the independent variables.

annual average daily traffic"). In equation 2, vehicle miles traveled is defined as daily vehicle miles traveled by county divided by county population.

In equation 3, county-level implied vehicle fuel efficiency is measured in miles per gallon and is computed by simply dividing the time series of vehicle miles traveled by the quantity of fuel sold, measured in gallons of gasoline equivalent. Recall that in equations 1 and 2, the dependent variables – fuel demand and vehicle miles traveled, respectively – contain annual observations of the average daily value per capita for each of the 67 Florida counties from 2001 to 2008. Therefore the units cancel, leaving a panel series of implied vehicle fuel efficiency.[72]

We compute a unique gasoline price series for each county assuming between-county variation results only from differing county tax rates. The average tax-inclusive price of gasoline in Florida is derived from the monthly PADD1c average price of "all grades all formulations" of retail gasoline, state tax rates and gasoline sales published by the U.S. Energy Information Administration.[73] Using Florida county tax rate history and gasoline sales statistics published by the Florida Energy and Climate Commission (2009), we construct a time series of the Florida's average county tax rate weighted by volume. The monthly county-level gasoline price is then computed as the Florida average monthly tax-inclusive price minus the state's average county tax rate plus the individual county tax rate. Instead of taking the simple average of the price of gasoline in each month, we weight the price in each month by the quantity of gasoline consumed to yield the average annual gasoline price.

---

[72] Whereas vehicle miles traveled are appropriately measured by county, fuel consumption measured by county sales records report consumption at the point of purchase rather than point of combustion. Thus implied fuel efficiency may be biased by out-of-county fuel purchases, particularly for commuters.

[73] PADD1c is composed of Florida, Georgia, North Carolina, South Carolina, Virginia and West Virginia.

The same procedure is performed to construct the average annual diesel price from the monthly PADD1c report of "No 2 diesel retail sales by all sellers" published by the U.S. Energy Information Administration (2009). The quantity of gasoline and diesel measured in gallons of gasoline equivalent are used to compute the weighted average fuel price for each county. Finally, this nominal fuel price series by county is deflated by the consumer price index for all urban consumers, all items obtained from the U.S. Department of Labor, Bureau of Labor Statistics (2009).

The measure of economic activity in our model is real gross sales of goods and services per resident. Data was compiled from the Florida Department of Revenue annual or semiannual reports to construct a single times-series of nominal gross sales by county. This data series was then divided by county population and deflated by the U.S. consumer price index for all urban consumers to yield the real gross sale of goods and services per resident in each county.[74, 75]

County-level average motor vehicle ownership is calculated by dividing motor vehicle registrations by county population. Statistics on total motor vehicle registrations by county were collected from the Florida Department of Highway Safety and Motor Vehicles. The data series "does not include mobile homes, trailers, vessels, dealer or

---

[74] The consumer price index for all urban consumers, all items is available from the U.S. Department of Labor, Bureau of Labor Statistics.

[75] Recall county gasoline and diesel fuel consumption were derived from county tax receipts. Because off-road vehicles used for farming, for example, are generally exempt from fuel taxes, our consumption data represents only on-road fuel demand. Similarly, vehicle miles traveled includes only public roads and our vehicle registration data is a proxy for the number on-road vehicles by county. Thus, or system of equations models on-road vehicle fuel demand, on-road vehicle miles traveled and implied on-road fuel efficiency in only the transportation sector.

transporter license plates, half-year truck/tractor registrations or permanent government license plates" and therefore is a reasonable proxy for on-road vehicles.[76]

Population density is calculated by dividing total county population by the county land area, measured in square miles. County population statistics were collected from the U.S. Census Bureau, Population Division (2009). The *2008 Florida Motor Gasoline and Diesel Fuel Report* contains land area statistics published by the *Florida Statistical Abstract 2007.*

e. Estimation Procedure

Recall the system of equations in (III-19) describing the dynamic models of vehicle fuel demand, vehicle miles traveled and fuel efficiency.

Equation 1:

$$\ln g_{i,t} = \mu_1 + \alpha_{1,i} + \beta_{1,1} \ln p_{i,t} + \beta_{1,2} \ln s_{i,t} + \beta_{1,3} \ln v_{i,t} + \beta_{1,4} \ln d_{i,t} + \left(1-\gamma_{1,1}\right) \ln g_{i,t-1} + u_{1,i,t}$$

Equation 2:

$$\ln m_{i,t} = \mu_2 + \alpha_{2,i} + \beta_{2,1} \ln p_{i,t} + \beta_{2,2} \ln s_{i,t} + \beta_{2,3} \ln v_{i,t} + \beta_{2,4} \ln d_{i,t} + \left(1-\gamma_{2,2}\right) \ln m_{i,t-1} + u_{2,i,t}$$

Equation 3:

$$\ln mpg_{i,t} = \mu_3 + \alpha_{3,i} + \beta_{3,1} \ln p_{i,t} + \beta_{3,2} \ln s_{i,t} + \beta_{3,3} \ln v_{i,t} + \beta_{3,4} \ln d_{i,t} + \left(1-\gamma_{3,2}\right) \ln m_{i,t-1} - \left(1-\gamma_{3,1}\right) \ln g_{i,t-1} + u_{3,i,t}$$

Before applying the constraints described in (III-20), the three models are estimated independently to test the specification of each model.

[76] Vehicle registration data is only available by fiscal year (starting in July and ending in June). We label the year the calendar year ending the fiscal year. For example, the fiscal year 2007-2008 corresponds with the year 2008 in our dataset.

Endogenous Regressors

The estimation of these three equations is complicated by the dynamic specification.[77] In particular, the lagged dependent variables are correlated with the county-specific effects causing the parameter estimates to be inconsistent. In Model 3, for example, implied fuel efficiency ($\ln mpg_{i,t}$) is a function of the county-specific effect ($\alpha_{3,i}$) by construction where Constraint A in (III-20) requires $\alpha_{3,i} = \alpha_{2,i} - \alpha_{1,i}$.[78] Thus, implied fuel efficiency is also a function of the county-specific effects in equations 1 and 2 ($\alpha_{1,i}$ and $\alpha_{2,i}$, respectively). Since the lagged dependent variables of equation 1 ($\ln g_{i,t-1}$) and equation 2 ($\ln m_{i,t-1}$) are also a function of the respective county-specific effect in each equation, the set of regressors in equation 3 is correlated with its combined residual ($\alpha_{2,i} - \alpha_{1,i} + u_{3,i,t}$). Therefore, standard OLS-based fixed or random effects estimation of equation 3 will yield biased and inconsistent parameter estimates. The same is true for equations 1 and 2.

To test the endogeneity of the lagged dependent variables ($\ln g_{i,t-1}$ and $\ln m_{i,t-1}$) each model is estimated independently using two-stage least squares (2SLS). The county-specific effects are assumed to be fixed.[79] In addition, the instruments for the lagged dependent variables are the current and lagged values of the regressors and are

---

[77] For a thorough discussion of panel data estimation see Baltagi (2008). For dynamic panel data estimation, see Greene (2003), Arellano and Bover (1995) and Ahn and Schmidt (1995) for example.

[78] The null hypothesis that $\alpha_{3,i} = \alpha_{2,i} - \alpha_{1,i}$ cannot be rejected in the unconstrained 2SLS-within estimation of the three equations. The results are presented in Table III-6.

[79] Below it is shown that the 2SLS-GLS estimator is rejected in favor of the 2SLS-within estimator in the unconstrained estimation procedure, suggesting the fixed effects specification is warranted. The results are presented in Table III-4.

identical in the estimation of each of the three models: $\ln p_{i,t}$, $\ln p_{i,t-1}$, $\ln s_{i,t}$, $\ln s_{i,t-1}$, $\ln v_{i,t}$, $\ln v_{i,t-1}$, $\ln d_{i,t}$, $\ln d_{i,t-1}$, $\ln g_{i,t-2}$ and $\ln m_{i,t-2}$.

A Durbin-Wu-Hausman (DWH) test is performed on the combined results of the three equations.[80] The DWH test is a Hausman-type test comparing the 2SLS-within parameter estimates to those of the OLS-within estimator. The covariance matrices for both 2SLS-within and the OLS-within estimator are based on the estimated error variances of the OLS-within estimator. The null hypothesis is that the OLS-within estimator of the equation is consistent. A rejection of the null hypothesis indicates the regressor (or set of regressors) being tested is correlated with the combined residual and should be treated as endogenous.

In the three equations of fuel demand, vehicle miles traveled and implied fuel efficiency, the lagged dependent variables ($\ln g_{i,t-1}$ and $\ln m_{i,t-1}$) appear a total of four times. Thus, the DWH test statistic is distributed as $\chi^2(4)$ with a critical value of 9.45 at the 5% level and 13.28 at the 1% level. The DWH test statistic for the exogeneity of the lagged dependent variables is 11.42, suggesting $\ln g_{i,t-1}$ and $\ln m_{i,t-1}$ are jointly correlated with the three combined error terms, $\alpha_{1,i} + u_{1,i,t}$, $\alpha_{2,i} + u_{2,i,t}$ and $\alpha_{3,i} + u_{3,i,t}$ at the 5% significance level. Therefore, the lagged dependent variables are treated as endogenous in the remainder of this paper. Nonetheless, the exogeneity of the remaining regressors is not assumed.

---

[80] Each equation is estimated independently without the system constraints; however, the DWH test is performed on the joint estimation results. A rejection of the null hypothesis indicates the lagged dependent variable is jointly endogenous in the system of fuel demand, vehicle miles traveled and implied fuel efficiency.

Table III-2: Durbin-Wu-Hausman test for the exogeneity of regressors

Specification:	**A**	**B**	**C**	**D**	**E**
Regressor(s):	$\ln m_{i,t-1}$, $\ln g_{i,t-1}$	$\ln p_{i,t}$	$\ln s_{i,t}$	$\ln v_{i,t}$	$\ln d_{i,t}$
Distribution:	$\chi^2(4)$	$\chi^2(3)$	$\chi^2(3)$	$\chi^2(3)$	$\chi^2(3)$
CV 5%	9.45	7.81	7.81	7.81	7.81
CV 1%	13.28	11.34	11.34	11.34	11.34
Eq. 1-3, jointly	11.42**	13.40***	0.88	4.49	7.56

The null hypothesis is that parameter estimates are consistent. Rejection implies the regressor is endogenous in the three models of fuel demand, vehicle miles traveled and fuel efficiency.
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

If Florida residents are strictly price-takers in the market for vehicle fuel, the assumption of price exogeneity would be warranted. However, a DWH test reveals the assumption yields inconsistent parameter estimates. In particular, in Specification B of Table III-2, price ($\ln p_{i,t}$) is removed from the set of exogenous regressors and added to the set of endogenous regressors which now becomes $\ln g_{i,t-1}$, $\ln m_{i,t-1}$ and $\ln p_{i,t}$. 2SLS-within is performed instrumenting for the three endogenous variables where appropriate in each model. The DWH test compares the parameter estimates of Specifications A and B. The test statistic, distributed as $\chi^2(3)$, is 13.40 and the 1% critical value is 11.34. Thus, the parameter estimates of Specification A are rejected as inconsistent in favor of Specification B. From here forward, the price of vehicle fuel is assumed to be correlated with the composed error terms and treated as endogenous in our system of vehicle fuel demand, vehicle miles traveled and fuel efficiency.[81]

[81] This result may follow from the construction of the price series used in the system of equations model of vehicle fuel demand, vehicle miles traveled and implied fuel efficiency. In particular, county-level prices are not available and we computed individual county-level price data assuming between county variation results only from differing county tax rates. However, it may be true that transportation costs vary by county or that fuel prices (pre and post-tax) may be correlated with tourism, causing some counties to have consistently higher fuel prices than others. In this example, high tourism and high tax rates would cause vehicle fuel prices to be correlated with the individual effect.

In Specification C, county-level sales of goods and services ($\ln s_{i,t}$) is removed from the set of exogenous regressors and added to the set of endogenous regressors which now becomes $\ln g_{i,t-1}$, $\ln m_{i,t-1}$, $\ln p_{i,t}$ and $\ln s_{i,t}$. 2SLS-within is performed instrumenting for the four endogenous variables. The DWH test compares the parameter estimates of Specifications B and C. The test statistic, distributed as $\chi^2(3)$, is 0.88, which does not exceed the critical value of 7.81 at the 5% level. Thus, the null hypothesis cannot be rejected, and county-level sales per capita is assumed to be exogenous.

Next, the endogeneity of vehicle ownership is tested. In Specification D, the share of vehicle ownership ($\ln v_{i,t}$) is added to the set of endogenous regressors which becomes $\ln g_{i,t-1}$, $\ln m_{i,t-1}$, $\ln p_{i,t}$ and $\ln v_{i,t}$. 2SLS-within is performed instrumenting for the four endogenous variables. The DWH test is performed comparing the parameter estimates of Specifications B and D. The test statistic, distributed as $\chi^2(3)$, is 4.49, which does not exceed the critical value of 7.81 at the 5% level. The null hypothesis cannot be rejected, and the share of vehicle ownership is assumed to be exogenous to our system of vehicle fuel demand, vehicle miles traveled and fuel efficiency.

In Specification E, population density ($\ln d_{i,t}$) is removed from the set of exogenous regressors and added to the set of endogenous regressors which now becomes $\ln g_{i,t-1}$, $\ln m_{i,t-1}$, $\ln p_{i,t}$ and $\ln d_{i,t}$. 2SLS-within is performed instrumenting for the four endogenous variables. The DWH test is performed comparing the parameter estimates of Specifications B and E. The test statistic, distributed as $\chi^2(3)$, is 7.56, which does not exceed the critical value of 7.81 at the 5% level. Therefore, the null hypothesis cannot be rejected, and population is assumed to be exogenous to our system of vehicle fuel demand, vehicle miles traveled and fuel efficiency.

County-Specific Effects

It remains to test the significance and nature of the county-specific effects in the individual equations for vehicle fuel demand, vehicle miles traveled and implied fuel efficiency. In the analysis that follows, we use 2SLS panel estimators where we instrument price ($\ln p_{i,t}$), lagged fuel demand ($\ln g_{i,t-1}$) and lagged vehicle miles traveled ($\ln m_{i,t-1}$) with the current and lagged values of the regressors: $\ln s_{i,t}$, $\ln s_{i,t-1}$, $\ln v_{i,t}$, $\ln v_{i,t-1}$, $\ln d_{i,t}$, $\ln d_{i,t-1}$, $\ln p_{i,t-1}$, $\ln g_{i,t-2}$ and $\ln m_{i,t-2}$.

First, we test the hypothesis that the county-specific effects are significant. The unrestricted model is the 2SLS-within estimator allowing for only county-specific heterogeneity in the intercept where the slopes are assumed to be common. The restricted model is the pooled-2SLS model with a common intercept and common slopes. There are $N-1$ restrictions and $N(T-1)-K$ degrees of freedom in each of the three equations. Wooldridge (1990), however, points out that the distribution of the F-statistic for the 2SLS estimator is unknown, even asymptotically. Thus, inferences made from the standard F-statistic can be misleading. Instead, Wooldridge suggests using the sum of squared residuals from the second-stage regression for both the restricted and unrestricted estimations in the numerator. The denominator remains the residuals from the 2SLS estimation.

The results of the test of the significance of the county-specific effects are presented in Table III-3. The 2SLS F-statistics of equations 1 and 2 are distributed as

$F(66,464)$ with a critical value of 1.33 at the 5% level and 1.50 at the 1% level. Due to the presence of an additional lagged dependent variable in the model of implied fuel efficiency, the equation 3 statistic is distributed as $F(66,463)$ with critical values of 1.33 and 1.50 at the 5% and 1% levels, respectively.[82] The null hypothesis that the county-specific effects are jointly equal to zero is rejected in equation 1 at the 5% significance level with an F-statistic equal to 1.40. The null hypothesis is easily rejected at the 1% level in the individual estimates of equations 2 and 3 with F-statistics of 6.48 and 6.43, respectively. Also, a test of the null hypothesis that the county-specific effects are jointly equal to zero in all three equations is rejected at the 1% level. The F-statistic, distributed as $F(198,1391)$, is equal to 4.77 which exceeds the critical value of 1.27 at the 1% level.

Table III-3: F-test for the significance of effects, unconstrained 2SLS-within

Equation:	Null Hypotheses:	Distribution:	CV 5%:	CV 1%:	F-stat:
Eq. 1: $\ln g_{i,t}$	$\alpha_{1,i} = 0\ \forall i = 1,\ldots,N$	$F(66,464)$	1.33	1.50	1.40**
Eq. 2: $\ln m_{i,t}$	$\alpha_{2,i} = 0\ \forall i = 1,\ldots,N$	$F(66,464)$	1.33	1.50	6.48***
Eq. 3: $\ln mpg_{i,t}$	$\alpha_{3,i} = 0\ \forall i = 1,\ldots,N$	$F(66,462)$	1.33	1.50	6.43***
Eq. 1-3, jointly	$\alpha_{j,i} = 0\ \forall j = 1,2,3\ \forall i = 1,\ldots,N$	$F(198,1391)$	1.19	1.27	4.77***

Under the null hypothesis, the county-specific effects are insignificant.
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

Finally, we test whether the county-specific effects are to be treated as fixed or random. The fixed effects estimator allows the effect to be correlated with the regressors and is always unbiased and consistent. However, if the unobserved county effects are strictly uncorrelated with the regressors, the random effects estimator is consistent and efficient since there are fewer parameters to estimate. Thus, we perform a Hausman test to compare the estimated coefficients. Under the null hypothesis, the effects are

[82] The critical values of the F-test of the significance of the county-specific effects in Model 3 are equal to those of Models 1 and 2 due to rounding.

uncorrelated with the regressors. Therefore, there will be no systematic difference between the fixed and random effects parameter estimates since each will converge to its true value.

In this Hausman test, 2SLS-within is the consistent estimator while 2SLS-GLS is the efficient estimator (and consistent under the null hypothesis). The test is distributed as $\chi^2(5)$ in equations 1 and 2 with a critical value of 15.09 at the 1%. In equation 3 the test is distributed as $\chi^2(6)$ with a critical value of 16.81 at the 1%. The results are presented in Table III-4. The 2SLS-GLS estimator is easily rejected at the 1% level in each of the three equations individually. The test statistics are equal to 39.42, 372.64, 388.80 in equations 1, 2 and 3, respectively.

Table III-4: Hausman test for random effects

Equation:	Distribution:	CV 5%:	CV 1%:	Test stat:
Eq. 1: $\ln g_{i,t}$	$\chi^2(5)$	11.07	15.09	39.42***
Eq. 2: $\ln m_{i,t}$	$\chi^2(5)$	11.07	15.09	372.64***
Eq. 3: $\ln mpg_{i,t}$	$\chi^2(6)$	12.59	16.81	3.88.80***
Eq. 1-3, jointly	$\chi^2(16)$	26.30	32.00	841.74***

Under the null hypothesis the 2SLS-GLS estimator is consistent.
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

Moreover, a Hausman test of the three equations estimated jointly is performed. The test statistic is distributed as $\chi^2(16)$ with a 1% critical value of 32.00. With a test statistic of 841.74, the 2SLS-GLS estimator is easily rejected in favor of the 2SLS-within estimator of the three jointly estimated system of equations for fuel demand, vehicle miles traveled and implied fuel efficiency. For the remainder of this paper, the effects are assumed to be significant, correlated with the regressors and are treated as fixed.

Table III-5: Unconstrained 2SLS-within parameter estimates

	Eq. 1: $\ln g_{i,t}$	**Eq. 2:** $\ln m_{i,t}$	**Eq. 3:** $\ln mpg_{i,t}$
Estimated short-run coefficients			
μ_j	0.2791 (0.2048)	3.3661*** (0.)	3.3659*** (0.2978)
$\beta_{j,1} \ln p_{i,t}$	-0.0385*** (0.0114)	-0.0021 (0.0127)	0.0416** (0.0164)
$\beta_{j,2} \ln s_{i,t}$	0.0483*** (0.0143)	0.0650*** (0.0151)	0.0128 (0.0190)
$\beta_{j,3} \ln v_{i,t}$	0.0701*** (0.0262)	-0.0389 (0.0286)	-0.0822** (0.0352)
$\beta_{j,4} \ln d_{i,t}$	-0.0938* (0.0542)	-0.3199*** (0.0563)	-0.1792** (0.0716)
$(1-\gamma_{j,1}) \ln g_{i,t-1}$	0.7021*** (0.0607)	--- ---	0.4801*** (0.0807)
$(1-\gamma_{j,2}) \ln m_{i,t-1}$	--- ---	0.3486*** (0.0348)	0.2160*** (0.0429)
Implied long-run coefficients			
c_j	0.9369	5.1676***	---
$b_{j,1} \ln p_{i,t}$	-0.1293**	-0.0032	---
$b_{j,2} \ln s_{i,t}$	0.1620***	0.0998***	---
$b_{j,3} \ln v_{i,t}$	0.2354**	-0.0597	---
$b_{j,4} \ln d_{i,t}$	-0.3149*	-0.4911***	---
Within R^2	0.4703	0.3639	0.3004
Overall R^2	0.8573	0.6733	0.5129

The subscript-*j* in the parameter notation refers to the model number. The long-run coefficients can not be recovered from the unconstrained 2SLS estimation of Model 3.
*, **, *** Denotes the statistic is significant at the 10%, 5% or 1% level, respectively.

Unconstrained Parameter Estimates

The unconstrained 2SLS-within parameter estimates of equations 1, 2 and 3 are presented in. The short-run parameter estimates (see (III-19)) are presented along with the standard errors, which are reported in parentheses. The implied long run coefficient estimates (see (III-21)) are presented together with the within equation and overall R^2. In general, all short and long-run parameter estimates significantly different from zero have the appropriate sign. In equation 1 describing fuel consumption, the short (and therefore long-run) coefficient on the price of fuel is negative and significant while the short (and long-run) coefficients on county-level sales, vehicle ownership and lagged fuel consumption are positive and significant. However, the short (and long-run) coefficient on density is significant at only the 10% level, but the negative sign is consistent with a review of the literature which suggests fuel consumption decreases with increasing density.

In equation 2 describing vehicle miles traveled, the short (and long-run) coefficients on the price of fuel and vehicle ownership are insignificant while the parameter estimates of county-level sales and lagged vehicle miles traveled are positive and significant. Finally, the coefficient on density is negative and significant.

In equation 3 describing implied vehicle fuel efficiency, the short-run coefficient estimate on county-level sales is insignificant. However, the short-run coefficients on price, lagged fuel consumption and lagged vehicle miles traveled are significant and positive, while for vehicle ownership and population density the coefficients are negative and significant.

Without Constraints B and C (see (III-20)), it cannot be guaranteed that $b_{3,k} = b_{2,k} - b_{1,k}$. Therefore, the long-run parameter estimates ($b_{3,k}$ for all $k = 1,...,4$) in equation 3 cannot be recovered. In particular, $b_{3,k} = \beta_{2,k} / \gamma_{3,2} - \beta_{1,k} / \gamma_{3,1}$ for all $k = 1,...,4$. While the single-equation estimation procedure does provide parameter estimates of $\gamma_{3,1}$ and $\gamma_{3,2}$, it does not provide estimates of $\beta_{1,k}$ or $\beta_{2,k}$. Thus, the long-run parameter estimates of equation 3 are not reported in

Table III-5.

The unconstrained 2SLS-within parameter estimates reported in

Table III-5 are used to test the 73 constraints described in (III-20). The results are presented in Table III-6. Constraint A is a set of 66 restrictions on the county-specific effects such that $\alpha_{2,\mathrm{i}} - \alpha_{1,i} = \alpha_{3,i}$ for each $i = 1,\ldots,66$. A single test of these restrictions is distributed as $F(67,1391)$. The critical value is equal to 1.31 at the 5% level. With an F-statistic of 0.14, the null hypothesis cannot be rejected. Therefore Constraint A is satisfied in the unconstrained 2SLS-within estimation procedure.

Table III-6: F-test of system constraints

Null Hypothesis:	Distribution:	CV 5%:	CV 1%:	F-stat:
Constraint A				
$\alpha_{2,\mathrm{i}} - \alpha_{1,i} = \alpha_{3,i}\ \forall i = 1,\ldots,N-1$	$F(66,1391)$	1.31	1.46	0.14
Constraint B				
$\mu_2 - \mu_1 = \mu_3$	$F(1,1391)$	3.85	6.65	0.41
Constraint C				
$\ln p_{i,t}: \beta_{2,1} - \beta_{1,1} = \beta_{3,1}$	$F(1,1391)$	3.85	6.65	0.05
$\ln s_{i,t}: \beta_{2,2} - \beta_{1,2} = \beta_{3,2}$	$F(1,1391)$	3.85	6.65	0.02
$\ln v_{i,t}: \beta_{2,3} - \beta_{1,3} = \beta_{3,3}$	$F(1,1391)$	3.85	6.65	0.26
$\ln d_{i,t}: \beta_{2,4} - \beta_{1,4} = \beta_{3,4}$	$F(1,1391)$	3.85	6.65	0.20
Constraint D				
$\ln g_{i,t-1}: \gamma_{1,1} = \gamma_{3,1}$	$F(1,1391)$	3.85	6.65	5.76**
$\ln m_{i,t-1}: \gamma_{2,2} = \gamma_{3,2}$	$F(1,1391)$	3.85	6.65	4.83**
Constraints A, B, C and D jointly	$F(73,1391)$	1.30	1.44	0.14

Under the null hypothesis each constraint is satisfied.
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

Constraint B requires $\mu_2 - \mu_1 = \mu_3$. The test of this restriction is distributed as $F(1,1391)$ with a critical value of 1.31 at the 5%. The F-statistic is equal to 0.41; therefore, the null hypothesis that $\mu_2 - \mu_1 = \mu_3$ cannot be rejected.

Constraint C is composed of four separate restrictions on the contemporaneous regressors. Specifically, $\beta_{2,k} - \beta_{1,k} = \beta_{3,k}$ for $k = 1,...,4$. The four tests are each distributed as $F(1,1391)$ with a 5% critical value equal to 3.85. The F-statistics are 0.05, 0.02, 0.26 and 0.20, respectively, for price, county-level sales, vehicle ownership and density. Thus Constraint C cannot be rejected in the unconstrained 2SLS-within estimation procedure.

Constraint D requires the dynamic adjustment of vehicle fuel demand to be equal in the models of vehicle fuel demand and implied fuel efficiency, i.e. $\gamma_{1,1} = \gamma_{3,1}$. Similarly, the dynamic adjustment of vehicle miles traveled must be equal in the models of vehicle miles traveled and implied fuel efficiency, i.e. $\gamma_{2,2} = \gamma_{3,2}$. Each of these constraints has one restriction. Thus the two F-tests are each distributed as $F(1,1390)$ with a critical value of 3.85 at the 5% level and 6.65 at the 1% level. The F-statistics are 5.76 and 4.83 for dynamic adjustments of vehicle fuel consumption and vehicle miles traveled, respectively. Therefore, the unconstrained 2SLS-within estimation procedure does not satisfy Constraint D at the 5 % level.

Finally, a test is performed on the null hypothesis that all 73 constraints described by A, B, C and D are jointly met. The F-statistic, distributed as $F(73,1390)$, is equal to 0.14, which does not exceed even the 5% critical value of 1.30. The null hypothesis cannot be rejected, and it is assumed that the unconstrained 2SLS-within estimation procedure satisfies the constraints of the system of vehicle fuel demand, vehicle miles traveled and implied fuel efficiency.

Constrained 2SLS Parameter Estimates

In order to recover the long-run parameter estimates of equation 3, the system of equations described in (III-19) is estimated via 2SLS-within including the 73 constraints defined in (III-20). F-tests confirm the county-specific effects remain significantly different than zero in the constrained 2SLS-within estimation procedure. The results are presented in Table III-7.

Table III-7: F-test for the significance of effects, constrained 2SLS-within

Equation:	Null Hypotheses:	Distribution:	CV 5%:	CV 1%:	F-stat:
Eq. 1: $\ln g_{i,t}$	$\alpha_{1,i} = 0\ \forall i = 1,\dots,N$	$F(66,1391)$	1.31	1.46	2.25***
Eq. 2: $\ln m_{i,t}$	$\alpha_{2,i} = 0\ \forall i = 1,\dots,N$	$F(66,1391)$	1.31	1.46	11.68***
Eq. 3: $\ln mpg_{i,t}$	$\alpha_{3,i} = 0\ \forall i = 1,\dots,N$	$F(66,1391)$	1.31	1.46	12.33***
Eq. 1-3, jointly	$\alpha_{j,i} = 0\ \forall j = 1,2,3\ \forall i = 1,\dots,N$	$F(132,1391)$	1.22	1.32	7.15***

Under the null hypothesis the county-specific effects are insignificant.
** Denotes the statistic is significant at the 5% level.
*** Denotes the statistic is significant at the 1% level.

In equations 1, 2 and 3, the test is distributed as $F(66,1391)$ with critical values equal to 1.31 and 1.46 at the 5% and 1% significance levels, respectively. The F-statistics are equal to 2.25, 11.68 and 12.33 respectively in equations 1, 2 and 3. The individual tests that the county-specific effects are jointly equal to zero in each of the three equations therefore is rejected. Due to the constraint that $\alpha_{2,i} - \alpha_{1,i} = \alpha_{3,i}$ for all $\forall i = 1,\dots,N-1$, the test on the system of equations has only 132 restrictions. Thus the F-statistic is distributed as $F(132,1391)$ with critical values equal to 1.22 and 1.32 at the 5% and 1% significance levels, respectively. The F-statistic is equal to 7.15, and the null hypothesis that the county-specific effects are jointly equal to zero in the system of all three equations is easily rejected.

The short-run parameter estimates are presented in Table III-8 along with the standard errors, reported in parentheses. The implied long-run parameter estimates are presented at the bottom of the table.

In general, all short and long-run parameter estimates significantly different than zero have the appropriate sign. Moreover, the constrained results are quite similar to the unconstrained estimation results, although several parameter estimates are more significant. For example, the short and long-run coefficients on population density in equation 1 are significant at the 10% level in the unconstrained case and at the 5% level in the constrained case. The short-run coefficients on price, vehicle ownership and density also become more significant in the model of fuel efficiency in equation 3.

In equation 1, $\gamma_{1,1} = 0.3731$ implies roughly 37% of the total adjustment to the long-run equilibrium demand for vehicle fuel occurs in the first year. In contrast, $\gamma_{2,2} = 0.7006$ implies roughly 70% of the total adjustment to the long-run equilibrium demand for vehicle miles traveled occurs in the first year. An F-test of the restriction that $\gamma_{1,1} = \gamma_{2,2}$ is distributed as $F(1,1391)$. With an F-statistic equal to 35.8, the hypothesis that the rate of adjustment to the long-run equilibrium demand for fuel is equal to the rate of adjustment to the long-run equilibrium demand for vehicle miles traveled is rejected. Thus, the equation for fuel efficiency implied by (III-16) is rejected in favor of the model implied by (III-15). Moreover, the results suggest that the demand for vehicle miles traveled responds more quickly to contemporaneous factors than the demand for vehicle fuel.

Table III-8: Constrained 2SLS-within parameter estimates

	Eq. 1: $\ln g_{i,t}$	**Eq. 2:** $\ln m_{i,t}$	**Eq. 3:** $\ln mpg_{i,t}$
Estimated short-run coefficients			
μ_j	0.3202* (0.1773)	3.5145*** (0.1992)	3.1943*** (0.2159)
$\beta_{j,1} \ln p_{i,t}$	-0.0327*** (0.0098)	0.0061 (0.0104)	0.0389*** (0.0118)
$\beta_{j,2} \ln s_{i,t}$	0.0524*** (0.0123)	0.0681*** (0.0126)	0.0157 (0.0140)
$\beta_{j,3} \ln v_{i,t}$	0.0676*** (0.0227)	-0.0317 (0.0238)	-0.0993*** (0.0260)
$\beta_{j,4} \ln d_{i,t}$	-0.1123** (0.0468)	-0.3227*** (0.0470)	-0.2104*** (0.0526)
$(1-\gamma_{j,1}) \ln g_{i,t-1}$	0.6269*** (0.0488)	--- ---	0.6263*** (0.0489)
$(1-\gamma_{j,2}) \ln m_{i,t-1}$	--- ---	0.2994*** (0.0269)	0.2994*** (0.0269)
Implied long-run coefficients			
m_j	0.8581	5.0163***	4.1581***
$b_{j,1} \ln p_{i,t}$	-0.0877***	0.0088	0.0965***
$b_{j,2} \ln s_{i,t}$	0.1405***	0.0972***	-0.0433
$b_{j,3} \ln v_{i,t}$	0.1812***	-0.0452	-0.2264***
$b_{j,4} \ln d_{i,t}$	-0.3009**	-0.4605***	-0.1596

The subscript-*j* in the parameter notation refers to the model number.
*, **, *** Denotes the statistic is significant at the 10%, 5% or 1% level, respectively.

Finally, we are able to separate the elasticity estimates of vehicle fuel demand into elasticity estimates of vehicle miles traveled and implied fuel efficiency. First, in the short-run (and long-run) the parameter estimate on the price of vehicle fuel ($\ln p_{i,t}$) is not significant in the equation for vehicle miles traveled, but it is positive and significant in the model of implied fuel efficiency. This evidence suggest that in the short-run drivers attempt to optimize engine performance by choosing less congested routes, performing routine maintenance or simply by driving at a lower speed at the cost of added time expense, rather than reducing overall vehicle miles traveled by carpooling or taking public transportation for example.

The long-run price elasticity of fuel demand is significant and equal to -0.0877, which can be disaggregated into the long-run price elasticity of vehicle miles traveled, 0.0088, and the long-run price elasticity of vehicle fuel efficiency, 0.0965. This is consistent with our findings in the short-run. However, roughly 60% of the improvement in vehicle fuel efficiency caused by increasing fuel prices occurs after the first year. This suggests that Florida drivers also respond to changes in price by improving the capital stock of vehicles. In other words, rising prices likely resulted in Florida residents purchasing more fuel efficient vehicles.

The impact of gross sales on vehicle fuel consumption, on the other hand, is dominated by a change in vehicle miles traveled rather than fuel efficiency. Specifically, the long-run elasticity of vehicle demand with respect to the sales of goods and services is significant and equal to 0.1405, which is composed of an elasticity of 0.0972 for vehicle miles traveled versus -0.0433 for vehicle fuel efficiency. This implies that an increase in economic activity results in an increase in vehicle miles traveled. Although the long-run coefficient on the sale of goods and services is not significant in the model of vehicle fuel efficiency in equation 3, the point estimate is negative. A loss of fuel efficiency might occur if increasing driving activity to acquire more goods causes reduced engine performance to do more frequent stops or added roadway congestions.

The long-run elasticity of vehicle fuel demand with respect to vehicle ownership is significant and equal to 0.1812, which is composed of an elasticity -0.0452 for vehicle miles traveled versus -0.2264 for vehicle fuel efficiency. Although one would anticipate the elasticity on vehicle miles traveled to be positive – implying that an increase in vehicle ownership leads to an increase in vehicle miles traveled – the statistic is not significantly greater than zero in our model. The result likely follows from the fact that vehicle ownership rates in Florida are already high. Specifically, there were 0.89 vehicles per Florida resident in 2008. However, the elasticity estimate of -0.2264 for vehicle fuel efficiency may indicate an increase in the share of residents owning vehicles leads to road congestion and poorer engine performance.

The short-run elasticity estimate of vehicle fuel demand with respect to density is significant and equal to -0.1123. This is consistent with both theoretical and empirical research which suggests that vehicle fuel demand and population density are negatively correlated. Given our system of equations, we are also able to decompose the impact of population density on vehicle fuel demand into the impact on vehicle miles traveled and vehicle fuel demand. In particular, the short-run elasticity estimate of vehicle miles traveled with respect to density is significant and equal to -0.3227, confirming vehicle miles traveled and population density are negatively correlated in the short-run in our sample. However, we find the short-run elasticity estimate of vehicle fuel efficiency with respect to population density to also be significant and equal to -0.2104. Thus, in the short-run an increase in population density decreases vehicle miles traveled and fuel demand, but a decrease in fuel efficiency – likely a result of increasing congestion – offsets 65% of the total vehicle fuel savings generated by driving few miles.

In the long-run, however, the elasticity estimates of fuel demand with respect to population density are not significant in every model. In particular, the long-run elasticity estimates of fuel demand and vehicle miles traveled with respect to population density are significant and equal to -0.3009 and -0.4605, respectively. This is consistent with the short-run results suggesting both fuel demand and vehicle miles traveled are negatively correlated with population density. However, the elasticity estimate of vehicle fuel efficiency is not significant, although it is negative and equal to -0.1596. Thus, in the long run as population increases, a decrease in fuel efficiency offsets about 35% of the fuel demand lost as vehicle miles traveled falls. This suggests that the impact of density on fuel efficiency (and therefore fuel demand) is muted in the long-run relative to the short-run. One explanation for this might be that poor engine performance is minimized in the long-run by infrastructure improvements designed to eliminate congestion.

f. Conclusions

In this paper, we developed a system of equations to model motor vehicle fuel consumption per resident, vehicle miles traveled per resident and implied vehicle fuel efficiency in the 67 counties in the State of Florida. With this system of equations we decomposed fuel demand elasticity estimates into elasticities of demand for vehicle miles traveled and motor vehicle fuel efficiency, and analyzed the way various factors influence fuel demand.

Particular attention is paid to the effects of vehicle fuel price, the gross sale of goods and services, vehicle ownership and population density on vehicle fuel demand, vehicle miles traveled and vehicle fuel efficiency. In summary, we find that an increase in the price of fuel results in a short and long-run decrease in fuel demand but not through a decrease in vehicle miles traveled. Instead we find evidence that a price increase leads to an increase in fuel efficiency. We also find that the value of goods sold is positively correlated with vehicle fuel demand as consumers likely increase vehicle miles traveled to acquire more goods. However, this has no significant impact on vehicle fuel efficiency. In addition, we find that an increase in the share of vehicle ownership does not result in an increase in vehicle miles traveled in our sample. Alternatively, a decrease in vehicle fuel efficiency causes vehicle fuel demand to increase. Finally, an increase in population density decreases vehicle miles traveled but the vehicle fuel savings is offset by 65% in the short-run, and 35% in the long-run, by a loss of fuel efficiency.

g. References

Ahn, S. and P. Schmidt (1995). "Efficient Estimation of Models for Dynamic Panel Data." *Journal of Econometrics*, 68(1): 5-27.

Anderson, W. P., P.S. Kanaroglou, and E.J. Miller (1996). "Urban Form, Energy and the Environment: A Review of Issues, Evidence and Policy." *Urban Studies*, 33 (1): 75-35.

Arellano, M (1987). "Computing Robust Standard Errors for Within-Groups Estimators." *Oxford Bulletin of Economics & Statistics*, 49(4): 431-434.

Arellano, M. and S. Bond (1991). "Some Tests of Specification for Panel Data: Monte Carol Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2), 277-297.

Archibald, R. and R. Gillingham (1981). "A Decomposition of the Price and Income Elasticities of the Consumer Demand for Gasoline." *Southern Economic Journal*, 47 (4): 1021-1031.

Banister, D. (1992). "Energy Use. Transportation and Settlement Pattern." In Breheny, M (ed.) *Sustainable Development and Urban Form*. Pion, London.

Balestra, P. and M. Nerlove (1966). "Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas." *Econometrica*, 34(3): 585-612.

Baltagi, B.H. (2008). *Econometric Analysis of Panel Data*. 4th ed. Chichester, UK: John Wiley & Sons Ltd.

Baltagi, B.H. and J.M. Griffin (1997). "Pooled Estimators vs. Their Heterogeneous Counterparts in the Context of Dynamic Demand for Gasoline." *Journal of Econometrics*, 77: 303-327.

Baltagi, B and Q. Li (1991). "A Joint Hypothesis Test for Serial Correlation for Serial Correlation and Random Individual Effects." *Statistics & Probability Letters*, 11: 277-280.

Bento, A.M., M.L. Cropper, A.M. Mobarak, and K. Vinha (2005). "The Effects of Urban Spatial Structure on Travel Demand in the United States." *The Review of Economics and Statistics*, 87(3): 466-478.

Bera, A., W. Sosa-Escudero and M. Yoon (2001). "Tests for the Error Component Model in the Presence of Local Misspecification." *Journal of Economics*, 101: 1-23.

Blundell, R. and S. Bond (1988). "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics,*84(1): 115-143.

Breusch, T.S. and A.R. Pagan (1980). "The Lagrange Multiplier Test and its Applications to Model Specification in the Presence of Local Misspecification." *Review of Economic Studies*, 47: 239-253.

Dahl, C. (1979). "Consumer Adjustment to a Gasoline Tax." *The Review of Economics and Statistics*, 61 (3): 427-432.

Dahl, C. (1982). "Do Gasoline Demand Elasticities Vary?" *Land Economics*, 58 (3): 373-382.

Dahl, C. and T. Sterner (1991). "Analyzing Gasoline Demand Elasticities: A Survey." *Energy Economics*, 13 (3): 203-210.

Dargay, J., D. Gately and M. Sommer (2007). "Vehicle Ownership and Income Growth, Worldwide: 1960-2010." *The Energy Journal*, 28 (4): 143-170.

Espey, M. (1998). "Gasoline Demand Revisited: An International Meta-Analysis of Elasticities." *Energy Economics*, 20 (3): 273-295.

Florida Department of Highway Safety and Motor Vehicles (2008). "10-Year Spreadsheet of Motor Vehicle Registrations for 2008." Available at www.flhsmv.gov/html/reports_and_statistics/10year/Registration.html.

Florida Department of Highway Safety and Motor Vehicles (2007). "10-Year Spreadsheet of Motor Vehicle Registrations for 2007." Available at www.flhsmv.gov/html/reports_and_statistics/10year/Registration.html.

Florida Department of Revenue, Office of Tax Research (2001-2008). "Sales Tax Return Collections by County (Form 9)." Available at http://dor.myflorida.com/dor/taxes/distributions.html

Florida Department of Transportation (2008). *2008 Source Book of Florida Highway Data.* Available at www.dot.state.fl.us/planning/statistics/sourcebook/.

Florida Department of Transportation, Transportation Statistics Office (2001-2008). "Public Road Mileage and Miles Traveled, 2001-2008." Available at www.dot.state.fl.us/planning/statistics/mileage%2Drpts/public.shtm

Florida Energy and Climate Commission (2009). *2008 Florida Motor Gasoline and Diesel Fuel Report.* Available at http://myfloridaclimate.com/climate_quick_links/florida_energy_climate_commission/policy_and_resources2/florida_motor_gasoline_and_diesel_fuel_reports.

Floyd, S.S., Ed. (2007). *Florida Statistical Abstract.* 42nd ed. Gainsville, FL: Bureau of Economic and Business Research, University of Florida.

Green, W.H. (2003). Econometric Analysis. 5th ed. New Jersey: Prentice-Hall.

Houthakker, H.A., P.K. Verleger Jr., and D. P. Sheehan (1974). "Dynamic Demand Analysis for Gasoline and Residential Electricity," *American Journal of Agricultural Economics*, 56 (2): 412-418.

Hughes, J.E., C.R. Knittel and D. Sperling (2006). "Evidence of a Shift in the Short-Run Elasticity of Gasoline Demand," NBER Working Paper, 12530.

Koyck, L.M. (1954). *Distributed Lags and Investment Analysis*. Amsterdam: North Holland Publishing.

Medlock III, K.B. and R. Soligo (2001). "Economic Development and End-Use Energy Demand," *The Energy Journal*, 22 (2): 77-105.

Medlock III, K.B. and R. Soligo (2002). "Automobile Ownership and Economic Development: Forecasting Passenger Vehicle Demand to the Year 2015," *Journal of Transport Economcs and Policy*, 36 (2): 163-188.

Mindali, O., A. Raveh, and I. Saloman (2004). "Urban Density and Energy Consumption: A New Look at Old Statistics," *Transportation Research Part A*, 38 (2): 143-162.

Newman, P. and J. Kenworthy (1989). *Cities and Automobile Dependence: A Source Book.* Brookfield, Vermont: Gower Publishing Company.

Office of Management and Budget (2009). "Update on the Statistical Area Definitions and Guidance on Their Uses," *OMB Bulletin No. 10-02*. Available at http://www.whitehouse.gov/omb/inforeg_statpolicy/.

Ostro, B.D. and J.L. Naroff (1980). "Decentralization and the Demand for Gasoline." Land Economics, 56 (2): 169-180.

Puller, S.L. and L.A. Greening (1999). "Household Adjustment to Gasoline Price Change: An Analysis Using 9 Years of US Survey Data," *Energy Economics*, 21 (1): 37-52.

Sanghi, A.K. (1976). "The Relationship between Population Density, Automobile Ownership and Automobile Use: Its Role in Transportation Planning," *Annals of Regional Science*, 10 (1): 118-127.

Shim, G.E., S.M. Rhee, and Ahn, K.H. (2006). "The Relationship between the Characteristics of Transportation Energy Consumption and Urban Form," *Annals of Regional Science*, 40 (20): 351-367.

Small, K.A. and Van Dender, A. (2007). "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect," *The Energy Journal*, 28 (1): 25-51.

Sosa-Escudero, W. and A.K. Bera (2001). "Specification Tests for Linear Panel Data Models," *Stata Technical Bulletin*, 61: 18-21.

StataCorp (2007). *Stata Longitudinal/Panel-Data Reference Manual: Release 10.* College Station, TX: StataCorp LP.

StataCorp (2007). *Stata Statistical Software: Release 9.* College Station, TX: StataCorp LP.

Sterner, T. and C. Dahl (1992). "Modelling Transport Fuel Demand," In Sterner, T., editor, *International Energy Economics*: 65-79. London: Chapman and Hall.

Stewart, C.T. and J.T. Bennet (1975). "Urban Size and Structure and Private Expenditure for Gasoline in Large Cities," *Land Economics*, 51 (4): 365-373.

Wooldride, J. (1990). "A Note on the Lagrange Multiplier and F-Statistics for Two Stage Least Squares Regressions." *Economics Letters*, 34: 151-155.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: The MIT Press.

Zellinsky, W. and D.F. Sly (1984). "Personal Gasoline Consumption, Population Patterns, and Metropolitan Structure: The United States, 1960-1970," *Annals of Association of American Geographers*, 74 (2): 257-278.

U.S. Census Bureau, Population Division (2009). "CO-EST2008-alldata: Annual Residen Population Estimates, Estimated Components of Resident Population Changes, and Rates of the Components of Resident Population Change for States and Counties: April 1, 2000 to July 1, 2008." Available at http://www.census.gov/popest/counties/counties.html.

U.S. Department of Labor, Bureau of Labor Statistics (2009). "Consumer Price Index for All Consumers: All Items." Available at http://research.stlouisfed.org/fred2/series/CPIAUCNS?cid=9.

U.S. Energy Information Administration (2009). "Table C1. Estimated Consumption of Vehicle Fuels in the United States, by Fuel Type, 2003-2007," *Alternatives to Traditional Transportation Fuels, 2007.* Available at http://www.eia.doe.gov/cneaf/alternate/page/atftables/attf_c1.html

U.S. Energy Information Administration (2009). "Table EN1. Federal and State Motor Fuels Taxes," *Petroleum Marketing Annual 2008.* Available at http://www.eia.doe.gov/oil_gas/petroleum/data_publications/petroleum_marketing_annual/pma.html.

U.S. Energy Information Administration (2009). "Table 5.13 Petroleum Consumption by Product by Sector, 1949-2008," *Annual Energy Review 2008*, Available at http://www.eia.doe.gov/emeu/aer/petro.html.

U.S. Energy Information Administration (2009). "Table 11. Transportation Sector Energy Consumption Estimates, Selected Years, 1960-2007, Florida," *State Energy Data System (SEDS).* Available at www.eia.doe.gov/emeu/states/state.html?q_state_a=fl&q_state=FLORIDA.

U.S. Energy Information Administration (2009). "Weekly Retail Gasoline and Diesel Prices," *Petroleum Navigator.* Available at http://tonto.eia.doe.gov/dnav/pet/pet_pri_gnd_dcus_r1z_m.htm